ISBN 978-623-325-629-6

Get Press Indonesia

# STATISTIK EKONOMI DAN BISNIS

Disusun oleh :

Pelliyezer Karo Karo, Anita Roosmawarni, Patrich Phill Edrich Papilaya
Toto Hermawan, Muhammad Zulkarnain, Roosganda Elizabeth
Joni Wilson Sitopu, Nadia Fazira, Asyidatur Rosmaniar, Derry Nugraha
Tedi Hartoyo dan Unang Atmaja, Dodi , Fitriana Dina Rizkina

# STATISTIK EKONOMI DAN BISNIS

**Pelliyezer Karo Karo**
**Anita Roosmawarni**
**Patrich Phill Edrich Papilaya**
**Toto Hermawan**
**Muhammad Zulkarnain**
**Roosganda Elizabeth**
**Joni Wilson Sitopu**
**Nadia Fazira**
**Asyidatur Rosmaniar**
**Derry Nugraha**
**Tedi Hartoyo dan Unang Atmaja**
**Dodi**
**Fitriana Dina Rizkina**

**GETPRESS INDONESIA**

**STATISTIK EKONOMI DAN BISNIS**

**Penulis :**
Pelliyezer Karo Karo
Anita Roosmawarni
Patrich Phill Edrich Papilaya
Toto Hermawan
Muhammad Zulkarnain
Roosganda Elizabeth
Joni Wilson Sitopu
Nadia Fazira
Asyidatur Rosmaniar
Derry Nugraha
Tedi Hartoyo dan Unang Atmaja
Dodi
Fitriana Dina Rizkina

**ISBN : 978-623-125-829-8**

**Editor :** Mila Sari, M. Si.
**Desain Sampul dan Tata Letak :** Tri Putri Wahyuni, S. Pd.

**PENERBIT : GET PRESS INDONESIA**
Anggota IKAPI No. 033/SBA/2022
Jl. Palarik RT 01 RW 06 Kelurahan Air Pacah
Kecamatan Koto Tangah Padang Sumatera Barat
website: www.getpress.co.id
email: adm.getpress@gmail.com

Cetakan pertama, Juni 2025

# KATA PENGANTAR

Puji dan syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa, yang telah memberikan rahmat dan karunia-Nya sehingga buku berjudul *Statistik Ekonomi dan Bisnis* ini dapat diselesaikan. Buku ini hadir untuk memberikan pemahaman terkait pengelolaan data. Isi buku ini mencakup pembahasan tentang pengantar statistik ekonomi dan bisnis, pengorganisasian dan penyajian data, ukuran pemusatan data, ukuran penyebaran data, ukuran letak data, konsep dasar probabilitas, distribusi probabilitas, metode pengambilan sampel, uji hipotesis, analisis varins (anova), korelasi dan regresi linear, regresi linear berganda, dan bilangan indeks.

Akhir kata, kami berharap buku ini dapat menjadi referensi yang bermanfaat bagi pembaca dalam pengumpulan data, pengolahan data, penyajian data, analisis data, dan interpretasi data.

Jakarta, Mei 2025

Penulis

## DAFTAR ISI

# DAFTAR GAMBAR

# DAFTAR TABLET

# BAB 1

# PENGANTAR STATISTIK EKONOMI DAN BISNIS

**Oleh Pelliyezer Karo Karo**

## 1.1 Konsep Dasar Statistik dalam Ekonomi dan Bisnis

### 1.1.1 Defenisi dan Ruang Lingkup Statistik

Statistik merupakan cabang ilmu yang berkaitan dengan lima tahapan pengelolaan data yaitu pengumpulan data, pengolahan data, penyajian data, analisis data, dan interpretasi data, dalam mendukung pengambilan keputusan yang lebih baik dengan berbasiskan pada data yang valid dan reliabel. Dalam konteks yang lebih luas, statistik tidak hanya mencakup pengelolaan data tetapi juga melibatkan perencanaan pengumpulan data dan pengambilan sampel yang representatif. Kata "statistik" dapat merujuk pada dua preferensi:

a. Statistik sebagai data dapat didefenisikan sebagai sekumpulan angka atau informasi yang diperoleh dari hasil pengukuran atau observasi. Contohnya adalah data tentang tingkat inflasi, jumlah pengangguran, atau pendapatan per kapita suatu negara.

b. Statistik sebagai ilmu merujuk pada metode atau teknik yang digunakan untuk mengumpulkan, mengolah, menganalisis, dan menyajikan data agar dapat digunakan dalam pengambilan keputusan. Kemudian dilanjutkan dengan menginter-pretastikan data, dimana tahapan ini dapat menjadi hal yang tidak terpisahkan menjadi satu kesatuan utuh ataupun menjadi pilihan tersendiri dengan beberapa limitasi. Sebab dalam tahapan ini, pemilihan fokus utama dalam menafsirkan dapat disesuaikan dengan skala prioritas kebutuhan.

Ruang lingkup statistik sangat luas dan mencakup berbagai bidang, baik dalam ilmu sosial, ekonomi, bisnis, sains, maupun teknologi. Berkaitan dengan aktivitas ekonomi dan bisnis, statistik berperan penting dalam memahami pola dan tren ekonomi, mengukur kinerja bisnis, serta membuat prediksi berdasarkan data historis. Berikut merupakan beberapa ruang lingkup utama statistik:

1. Pengumpulan Data

   Sebelum mengumpulkan data, terdapat dua kegiatan perencanaan yang penting diterapkan agar data yang dikumpulkan relevan dan representatif dengan kebutuhan informasi. Pertama, menetapkan besaran data dan kedua, mengidentifikasi karakteristik data yang sesuai dengan tingkat kedalaman informasi yang dibutuhkan. Setelah kedua hal tersebut ditetapkan, kemudian dilakukan proses mengumpulkan data. Dalam mengumpulkan data, metode yang efektif perlu dipilih dengan tepat agar prosesnya menjadi efektif dan efisien. Metode pengumpulan data yang umum digunakan antara lain observasi, survei atau kuesioner, wawancara, eksperimen atau studi pustaka. Pemilihan metode tentunya harus merujuk pada karakteristik data, seperti yang disampaikan dalam faktor

perencanaan sebelumnya. Pengumpulan data yang berhubungan dengan ekonomi dan bisnis dapat berupa preferensi konsumen, sensus penduduk, data transaksi pembelian dan penjualan dan lainnya.

2. Pengolahan Data

   Dalam melakukan pengolahan data, terdapat beberapa rangkaian proses mengorganisasikan data seperti data cleaning atau membersihkan data dari residunya, mengklasifikasikan atau mengelompokkan data, dan melakukan transformasi atau normalisasi data. Rangkaian tersebut krusial untuk dilakukan agar menghasilkan data yang akurat dan memiliki varians yang kecil, sehingga memberikan kemudahan dalam tahapan berikutnya. Bagi sektor ekonomi dan bisnis, dengan menjalankan rangkaian tersebut dengan tepat akan memberikan informasi yang tidak bias terkait pasar, konsumen, produk dan jasa, hingga pesaing. Pilihan pengambilan keputusan akan berbasiskan data yang valid, reliabel dan representatif terhadap kebutuhan ekonomi maupun bisnis.

   Ketiga rangkaian tersebut memiliki fungsi masing-masing yang saling berhubungan, jika dijalankan secara terstruktur akan mampu meningkatkan efisiensi bagi tahapan berikutnya yaitu analisis data.

   a. Membersihkan data dilakukan untuk menghapus atau memperbaiki datum atau data yang tidak valid, duplikat atau hilang dari kelompoknya. Melalui proses ini, akan mampu mendeteksi kesalahan input, kemungkinan datum yang belum lengkap dan data yang tidak relevan.

b. Mengklasifikasikan data berdasarkan kategori tertentu dilakukan agar mempermudah proses identifikasi dan analisis. Dengan meng-elompokkan data, memberikan peluang analisis yang lebih spesifik.
c. Melakukan transformasi dimaksudkan agar seluruh data yang dikumpulkan memiliki satuan yang sejenis sehingga mudah untuk diproses dan dibandingkan. Sedangkan, menormalisasikan data bertujuan meninjau seberapa luas data berada dalam rentang integritas atau standar kepercayaan yang ditetapkan. Proses ini dapat diterapkan dengan mengidentifikasi pengulangan data atau redudanasi dan memastikan kumpulan data terhubung secara konsisten,

3. Analisis Data

Analisis data secara statistik merupakan sebuah proses sistematis berdasarkan standar dan aturan ilmu statistik yang digunakan untuk mengelola data mentah (yang telah melalui tahapan sebelumnya) agar menjadi kumpulan informasi yang dapat berguna dalam pengambilan keputusan. Tahapan ini dapat digunakan untuk mempelajari banyak hal sesuai dengan tujuan utama, seperti mengidentifikasi hubungan, pengaruh, pola atau tren, prediksi, mediasi, derajat, dampak dan lain sebagainya. Tentunya, tujuan yang diinginkan harus menggunakan metode analisis yang sesuai sehingga informasi yang diperoleh relevan dengan kebutuhan. Dalam ekonomi dan bisnis, misalnya memahami tren pasar, mengukur kinerja iklan, mengukur dampak kebijakan ekonomi, mengevaluasi peraturan yang berkaitan dengan faktor ekonomi, mengidentifikasi kekuatan relasi antar faktor keputusan konsumen dan hal lainnya.

Terdapat dua proses utama dalam tahapan analisis data yaitu menentukan tujuan melakukan analisis dan memilih metode analisis yang sesuai untuk dijalankan.

a. Menentukan tujuan analisis

   Tujuan dalam melakukan analisis sebaiknya berdasarkan pada identifikasi permasalahan yang terjadi atau pilihan yang ingin diatasi. Dengan berbasis hal tersebut, maka variabel yang akan dipilih atau dipergunakan menjadi lebih jelas. Misalnya, jumlah pelanggan, harga produk, tingkat permintaan, tingkat penawaran, tingkat kepuasan konsumen, preferensi konsumen, tingkat loyalitas konsumen, *brand awareness, brand equity* dan lain sebagainya yang relevan dengan kebutuhan.

b. Memilih metode analisis

   Metode analisis yang dipilih berdasarkan tujuan analisis yang diinginkan. Selain itu, tingkat kedalaman analisis juga menyesuaikan dengan jenis data. Secara umum, analisis data secara statistik terbagi menjadi dua model, yaitu analisis statistik deskriptif dan analisis statistik inferensial. Analisis deskriptif bertujuan untuk memberikan gambaran data dengan ringkas yang berasal dari karakteristik populasi yang diamati sedangkan analisis inferensial digunakan untuk mengambil kesimpulan atau membangun prediksi yang berasal dari karakteristik sampel sebagai perwakilan populasi yang diamati.

   - Analisis statistik deskriptif terdiri atas teknik ukuran pemusatan data dan teknik ukuran penyebaran data. Ukuran

pemusatan data merupakan teknik yang dipergunakan dalam menggambarkan pusat dari kumpulan data. Ukuran penyebaran data merupakan teknik yang dipergunakan dalam menggambarkan jarak penyimpangan nilai dari kumpulan data.

- Analisis statistik inferensial dipahami sebagai metode yang digunakan untuk menguji hipotesis. Terdapat empat teknik analisis yang umum dipergunakan dalam metode ini, yaitu analisis korelasi, analisis regresi, analisis varian dan uji chi-square.

4. Penyajian Data

Hasil dari pengolahan data dan analisis data dapat divisualisasikan kedalam berbagai format dengan tujuan informasi yang disampaikan dapat lebih mudah dipahami. Secara umum, terdapat tiga format penyajian data yang sering dipergunakan, termasuk juga dalam ruang ekonomi dan bisnis. Format tersebut mencakup mengorganisir data dalam bentuk tabel, grafik atau diagram dan narasi atau teks.

a. Format tabel digunakan untuk menyajikan data secara terstruktur kedalam bentuk kolom dan baris. Dengan format ini, melakukan perbandingan data antar kategori menjadi lebih mudah. Data dengan jumlah atau nilai yang besar juga cocok ditanyangkan dalam format tabel.
b. Format grafik atau diagram digunakan untuk menyajikan data dengan membentuk pola dan tren, baik inter kelompok data maupun intra kelompok data. Jenis grafik atau diagram yang umum dipergunakan antara lain bar chart atau diagram batang, line chart atau diagram garis,

pie char atau diagram lingkaran, dan histogram atau grafik distribusi.

c. Format narasi digunakan untuk menyajikan data dalam bentuk kalimat dengan tujuan informasi yang ingin disampaikan menjadi lebih mudah dipahami oleh pengguna. Format ini biasanya menjadi pelengkap bagi dua format lainnya dan sering dipergunakan dalam naskah akademik ataupun laporan bisnis yang membutuhkan penjelasan dalam konteks yang lebih kecil atau lebih mengerucut kepada satu informasi tertentu.

5. Interpretasi Data

Setelah melakukan analisis data dan memberikan hasil, kemudian menyajikan data dari hasil analisis tersebut, maka proses interpretasi data menjadi tahapan terakhir. Proses ini merupakan rangkaian dalam menarik kesimpulan dari hasil analisis data yang dibantu secara visual melalui penyajian data. Hasil interpretasi berdasarkan temuan statistik dapat berupa memberikan rekomendasi, menyimpulkan hubungan antar variabel, menemukan pola atau tren yang terbentuk, membangun prediksi, mengkuantifikasi dampak hingga memberikan hasil evaluasi. Dalam melakukan interpretasi, terdapat tiga faktor umum yang menyebabkan kesalahan atau mis-interpretasi terhadap hasil analisis data.

a. Variabel lain terabaikan

Dalam mengelola data secara statistik tentunya terdapat batasan baik dalam seluruh tahapan statistik ataupun batasan yang dibangun diawal sesuai tujuan pengelolaan datanya. Batasan ini merupakan keterwakilan terhadap kemungkinan variabel lain yang berhubugan dengan fokus pengelolaan data

yang sedang terjadi. Oleh sebab itu, ketika melakukan interpretasi data, khususnya dengan metode statistik inferensial, mempertimbangkan adanya irisan dengan variabel lain sebaiknya disampaikan dalam hasil interpretasi data. Hal ini penting untuk selalu disampaikan agar mengurangi bias dari kesimpulan yang disampaikan. Misalnya, hasil analisis data menunjukkan bahwa harga menjadi pertimbangan bagi keputusan konsumen, bukan berarti seluruh kelompok konsumen tidak memiliki faktor lain yang menentukan keputusan pembeliannya.

b. Data tidak representatif

Melakukan generalisasi melalui sampel yang dianggap telah mewakili populasi merupakan tantangan tersendiri dalam pengelolaan data secara statistik. Apabila data yang dikumpulkan tidak mampu mencakup seluruh kelompok atau karakteristik yang terbangun dalam populasi maka jumlah sampel yang kecil menjadi tidak representatif bagi tujuan utama. Hal ini menjadi penyebab bias terbesar dalam menghasilkan kesimpulan. Misalnya, survei kepuasan konsumen yang dilakukan pada satu atau beberapa outlet, belum tentu dapat menjadi keterwakilan konsumen secara umum terhadap produk yang sedang dievaluasi.

c. Varian data tinggi

Varian data menunjukkan lebar distribusi data yang sedang dikelola. Semakin tinggi varian data mengindikasikan bahwa data memiliki rentang yang cukup besar dan antar data terdistribusi cukup jauh. Maknanya, kualitas data diragukan sehingga tidak dapat

dipergunakan sebagai acuan secara umum. Oleh sebab itu, menggunakan nilai mean atau rata-rata dengan tidak mempertimbangkan nilai standar deviasi akan menghasilkan kesimpulan yang bisa jadi tidak sesuai dengan kondisi fakta dan hal ini dapat disebut dengan rekomendasi menyesatkan. Misalnya, pengelolaan data tentang tingkat pendapatan masyarakat. Jika ditemukan rata-rata tingkat pendapatan masyarakat berada pada angka empat juta rupiah, akan tetapi varian data juga menunjukkan bahwa terdapat masyrakat dengan tingkat pendapatan terendah dibawah satu juta rupiah dan tingkat pendapatan tertinggi pada angka dua puluh juta rupiah, maka nilai rata-rata pendapatan masyarakat tidak tepat untuk dijadikan rujukan dalam distribusi tingkat pendapatan masyarakat.

### 1.1.2 Peran Statistik dalam Ekonomi dan Bisnis

Statistik memiliki peran penting dalam berbagai aspek, termasuk didalamnya ekonomi dan bisnis. Dengan menggunakan statistik, proses pengambilan keputusan menjadi lebih relevan dan *representative* sebab dibangun berbasis data yang telah valid dan reliabel. Statistik membantu menganalisis fenomena ekonomi dan bisnis secara objektif untuk memberikan informasi yang lebih akurat dan personal dalam meningkatkan efektifitas dan efisiensi dalam menjalankan operasional perusahaan.

1. Peran statistik dalam ekonomi

   Dalam bidang ekonomi, statistik dapat digunakan untuk mengukur, menganalisis hingga meramalkan kondisi ekonomi suatu wilayah atau

negara. Berikut beberapa peran yang menggunakan statistik sebagai penguatan data bidang ekonomi.

a. Pertumbuhan ekonomi

   Statistik dapat digunakan untuk mengukur kondisi neraca nasional atau kondisi nilai pasar suatu wilayah atau negara, melalui analisis terhadap pergerakan Produk Domestik Bruto (PDB) atau yang dikenal secara internasional sebagai *Gross Domestic Product* (GDP). Selain hal tersebut, statistik juga dapat membantu menganalisis faktor yang berkaitan dengan indikator pertumbuhan ekonomi lainnya seperti tingkat konsumsi, investasi, dan nilai ekspor impor.

b. Tingkat inflasi berbasis harga pasar

   Dalam mengukur tingkat inflasi pada periode tertentu, statistik menggunakan angka Indeks Harga Konsumen (IHK), dengan menghitung perubahan atau pertumbuhan harga barang dan jasa. Hasil monitoring harga pasar secara statistik dapat membantu pemerintah dan bank sentral dalam menentukan kebijakan moneter dan fiskal pada periode berjalan atau mempersiapkan kebijakan untuk periode mendatang, yang berbasis pada pengelolaan data historis.

c. Perencanaan pembangunan ekonomi

   Serupa dengan penggunaan statistik bagi pemantuan inflasi, pemerintah juga dapat menggunakan data statistik dalam ruang lingkup lebih besar untuk menyusun kebijakan ekonomi jangka pendek maupun jangka panjang. Dalam konteks pajak, misalnya perhitungan potensi peningkatan nilai pajak digunakan untuk menghasilkan kebijakan insentif pajak. Dalam konteks pariwisata,

misalnya, peningkatan proporsi tingkat kunjungan wisatawan dari beberapa negara potensial digunakan untuk menghasilkan kebijakan bebas Visa kunjungan bagi sejumlah negara tertentu.

d. Peramalan ekonomi

Selain digunakan untuk menghasilkan rekomendasi bagi periode terkini, statistik juga dapat digunakan untuk memprediksi kemungkinan situasi dan kondisi ekonomi di masa depan, berdasarkan tren data historis. Fungsi ini dapat digunakan pemerintah dan perusahaan dalam menentukan strategi yang paling tepat. Peramalan perkembangan ekonomi erat kaitannya dengan strategi investasi, kemitraan eskpor impor antar negara, kebijakan hilirisasi sumber daya alam dan berbagai hal lainnya yang mendukung pertumbuhan ekonomi.

e. Dinamika ketenagakerjaan

Dalam perjalanannya, salah satu tantangan pertumbuhan ekonomi adalah tingkat pengangguran. Statistik mengambil peran dalam mengukur seluruh faktor yang berkaitan dengan angkatan kerja, mulai dari kesempatan kerja, pendidikan tenaga kerja, kompetensi tenaga kerja, besaran upah kerja, struktur lapangan kerja dan hal terkait lainnya. Analisis terhadap data tersebut membantu pemerintah dalam merancang, mengembangkan hingga mengevaluasi seluruh peraturan dan kebijakan yang berkaitan dengan ketenagakerjaan.

2. Peran statistik dalam bisnis

   Sejalan dengan bidang ekonomi, dunia bisnis juga membutuhkan peran serta statistik untuk memberikan pertimbangan mendasar dan relevan, mulai dari yang berhubungan dengan konsumen hingga pesaing. Berikut beberapa peran statistik dalam penguatan bidang bisnis.

   a. Perilaku konsumen

      Analisis pasar selalu berkaitan antara perusahaan dengan konsumen. Dalam area konsumen, statistik berperan untuk menemukenali pola pembelian, mengidentifikasi target pasar paling potensial, memahami preferensi konsumsi dan hal lainnya. Misalkan, dilakukan survey kepuasan pelanggan dengan output hasil berupa persepsi konsumen terhadap produk yang dipilih. Melalui hasil analisis dan rekomendasi, perusahaan akan mampu meningkatkan kualitas produk atau jasa yang dihasilkan.

   b. Tren permintaan dan penjualan

      Data historis dari kelompok data permintaan dan kelompok data penjualan tidak hanya berguna untuk mengidentifikasi hubungan atau faktor pembentuknya. Akan tetapi dapat ditingkatkan menjadi rangkaian data terstruktur berupa tren permintaan dan tren penjualan. Melalui tren ini, perusahaan mampu melakukan perkiraan terhadap kebutuhan produk atau jasa di periode berikutnya. Sebagai contoh, statistik tren permintaan bagi perusahaan ritel menjadi referensi penting untuk menentukan rentang waktu terbaik kapan menambah ragam produk atau meningkatkan jumlah stok.

c. Efisiensi operasional

Dalam konteks perusahaan yang memiliki rangkaian proses produksi yang panjang, statistik dapat berperan menghitung alur kerja dan durasi dari setiap tahapan proses. Setiap tahapan dapat dianalisis sesuai kebutuhan perusahaan, apakah berbasis pada perhitungan beban kerja maksimum, jam kerja tenaga kerja, distribusi bahan baku dan hal lainnya. Melalui basis perhitungan tersebut, statistik mampu menemukan tahapan atau area yang berpotensi untuk dioptimalkan. Misalnya, dalam rangka meningkatkan produktivitas tenaga kerja pabrik furniture, maka dilakukan analisis waktu produksi setiap tahapan proses dan ditemukan bahwa terdapat dua tahapan yang dapat dimulai secara bersamaan dengan menerapkan sistem kontrol.

d. Pengendalian kualitas

Menjaga kualitas produk sangat perlu dilakukan untuk meminimalisir distribusi produk cacat. Semakin rendah jumlah produk cacat yang diterima konsumen, tingkat kepercayaan konsumen akan tetap terjaga dan berpotensi meningkat. Tentunya, tingginya jumlah produksi akan membuat *quality control* tidak mungkin hanya dilakukan oleh satu atau beberapa tenaga kerja. Menerapkan *Statistical Process Control* (SPC) dapat menjadi jawaban dalam kondisi tersebut. Untuk jumlah produksi yang tidak terlalu tinggi, perusahaan dapat melakukan kontrol kualitas dengan menerapkan teknik sampling statistika yang sesuai dengan karakteristik jenis produksi.

e. Risiko investasi

Tentunya dalam setiap jenis bisnis, risiko berjalan berdampingan dan tidak dapat dikesampingkan begitu saja tetapi tetap dapat diminimalisir dampaknya maupun frekuensinya melalui manajemen risiko. Risiko yang hadir dalam waktu yang bersamaan dengan prosesnya dan cenderung berdampak besar dan realtime dapat ditemukan pada sektor keuangan seperti saham, perbankan dan investasi. Sektor ini membutuhkan data dan informasi akurat di setiap waktu operasionalnya. Statistik berperan dalam penyediaan data dan pengolahan data untuk menghasilkan informasi terkini. Selain itu, analisis prediksi yang akurat dengan tingkat kepercayaan sangat tinggi juga diperlukan untuk memastikan perolehan nilai imbal balik yang positif bagi perusahaan maupun konsumen.

## 1.2 Jenis Statistik

Statistik dapat diklasifikan sesuai dengan kriteria kebutuhannya. Dua kriteria mendasar yang umum digunakan seperti tujuan analisis dan sifat data yang dipergunakan. Berdasarkan tujuan analisis, statistik terbagi menjadi dua yaitu statistik deskriptif dan statistik inferensial.

### 1.2.1 Statistik Deskriptif

Statistik deskriptif digunakan dalam kebutuhan pengelolaan data dengan sederhana. Jenis statistik ini menggunakan seluruh data yang dikumpulkan sebagai populasi pengelolaan, bukan sebagai perwakilan populasi. Statistik deskriptif tidak dapat digunakan

untuk menarik kesimpulan sebab bukan bertujuan untuk melakukan generalisasi kesimpulan. Tujuan utamanya fokus pada menyederhanakan raw data agar lebih mudah dimengerti oleh pengguna, melalui analisis pemuasatan data dan penyebaran data.

Tiga ukuran pemusatan data yang paling sering digunakan yaitu mean, median dan mode.

1. Mean merupakan nilai rata-rata dari kumpulan data.
2. Median merupakan nilai tengah dari kumpulan data setelah diurutkan, baik dari kecil ke besar atau sebaliknya.
3. Mode merupakan data dengan frekuensi tertinggi atau sering dipahami sebagai data yang paling sering muncul dari kumpulan data.

Tiga ukuran penyebaran data yang umum digunakan yaitu *range*, *variance* dan *standard deviation*.

1. *Range* merupakan rentang data, dipereloleh dengan menghitung selisih antara nilai maksimum dengan nilai minimum.
2. *Variance* merupakan keragaman data yang mengukur besaran jarak dari seluruh data tersebar dibandingkan dengan nilai rata-rata.
3. *Standard deviation* merupakan simpangan data yang menunjukkan seberapa besar adanya variasi dari kumpulan data.

### 1.2.2 Statistik Inferensial

Statistik inferensial dibutuhkan ketika pengelolaan data lebih kompleks, luas dan membutuhkan rekomendasi. Data dalam jenis statistik ini adalah data perwakilan dari populasi atau sering disebut dengan sampel. Oleh sebab itu, teknik sampling menjadi salah satu pilar penting agar kelompok data yang dikumpulkan mampu memenuhi sifat representatif, baik

secara jumlah, distribusi dan karakteristik data sebagai keterwakilan populasi. Secara umum, statistik inferensial dipergunakan untuk tujuan menarik kesimpulan sebagai bentuk generalisasi bagi populasi atau memberikan prediksi bagi populasi berikutnya. Dengan kata lain, pengujian hipotesis hanya dapat dilakukan dengan menerapkan statistik inferensial.

Berikut merupakan pengujian atau analisis yang sering diterapkan dalam berbagai bidang, termasuk ekonomi dan bisnis.

1. Analisis korelasi bertujuan untuk memahami sejauh mana dua variabel atau dua kelompok data saling bergantung. Dalam analisis ini, sering juga disebut dengan tingkat keeratan hubungan. Output tingkat keeratan antar variabel terbagi menjadi korelasi kuat, korelasi lemah dan tidak memiliki korelasi. Korelasi juga dapat bernilai negatif atau positif tergantung arah hubungan variabelnya. Misalnya, perusahaan sedang mengevaluasi bagaimana hubungan antara program membership dengan loyalitas konsumen. Jika ditemukan nilai korelasi positif dan lebih besar dari 0,8 maka perusahaan dapat menyimpulan bahwasanya program membership yang dijalankan memiliki hubungan yang kuat dan berbanding lurus dengan adanya loyalitas konsumen. Apabila dibutuhkan pembuktian lebih lanjut perihal apakah program membership yang menyebabkan peningkatan loyalitas konsumen dan seberapa besar pengaruh programnya terhadap loyalitas, maka dapat dilanjutkan dengan menggunakan analisis regresi.
2. Analisis varians bertujuan untuk menguji dan menganalisis signifikansi perbedaan antara dua atau lebih kumpulan data melalui nilai mean. Analisis ini juga lebih dikenal dengan akronim ANOVA (*analysis of variance*). Perbedaan yang diuji

melalui analisis ini dapat berupa perbedaan antar kelompok sampel maupun perbedaan antar perlakuan atau intervensi. Secara umum, terdapat tiga jenis metode pada analisis ini yaitu *one-way* ANOVA, *two-way* ANOVA dan *repeated measures* ANOVA.

a. Metode *one-way* digunakan untuk membandingkan mean antara tiga atau lebih kelompok data yang tidak saling berkaitan. Sebagai contoh, dilakukan penelitian untuk mengetahui apakah terdapat perbedaan yang signifikan dari tingkat daya beli anatara beberapa kelompok masyarakat dengan proritas kualitas hidup berbeda.
b. Metode *two-way* digunakan untuk membandingkan pengaruh antara dua atau lebih variabel terhadap hasil pengamatan pada kelompok sampel berbeda. Sebagai contoh, penelitian tentang pengaruh besar upah dan jenjang karir terhadap loyalitas tenaga kerja pada beberapa kelompok perusahaan.
c. Metode *repeated* digunakan untuk mengukur output yang sama pada kelompok yang sama pada waktu yang berbeda. Sebagai contoh, penelitian untuk mengidentifikasi perubahan dari standar hidup layak masyarakat pada masa sebelum, selama dan setelah pemberian subsidi penguatan ekonomi.

3. Analisis regresi linier bertujuan untuk memahami hubungan antara dua jenis variabel yang sedang diteliti. Salah satu jenis variabelnya merupakan variabel dependen atau variabel terikat dan jenis variabel lainnya adalah variabel independen atau variabel bebas. Variabel independent dapat berjumlah satu variabel atau lebih dari satu variabel. Jika variabel independent berjumlah satu

maka analisis yang digunakan adalah regresi linier sederhana sedangkan variabel independent yang berjumlah lebih dari satu akan menggunakan analisis regresi linier berganda.

Misalnya, analisis regresi linier berganda digunakan untuk melihat pengaruh antara harga produk terhadap keputusan pembelian. Sedangkan, analisis regresi linier berganda digunakan untuk memahami pengaruh antara harga produk, kualitas pelayanan, dan promosi terhadap keputusan konsumen untuk membeli.

4. Analisis uji t (t-*test*) digunakan untuk membandingkan apakah terjadi perbedaan signifikan antara satu atau dua kelompok pengamatan dengan nilai statistik yang dipersyaratkan. Pengujian dalam analisis ini terbagi *one sample t-test* dan *two sample t-test* yang terbagi menjadi *independent sample* dan *paired sample*. Pengujian *independent sample t-test* diterapkan apabila digunakan untuk menguji dua kelompok data yang bebas atau tidak memiliki hubungan, sedangkan *paired sample t-test* dibutuhkan apabila digunakan untuk menguji dua kelompok data yang saling berhubungan atau memiliki korelasi. Misalnya, pemerintah melakukan evaluasi dampak dari menerapkan kebijakan insentif pajak. Maka digunakan uji t sampel berpasangan (*paired*) untuk menghitung besaran t-hitung antara data capaian pajak reguler dengan kelompok data capaian setelah dilakukan intervensi melalui penerapan insentif.

### 1.2.3 Statistik Parametrik

Selain berdasarkan pada tujuan analisis dari pengelolaan data secara statistik, statistik juga dibagi menjadi dua jenis berdasarkan sifat data yang digunakan. Dua jenis tersebut yaitu statistik parametrik dan statistik non parametrik.

Jenis statistik parametrik umumnya digunakan untuk jenis data yang menggunakan skala rasio dan skala interval. Selain itu, statistik ini juga dapat diterapkan apabila telah memenuhi beberapa asumsi tertentu seperti asumsi homegenitas data melalui nilai varians atau asumsi data berdistribusi normal. Dalam prakteknya, beberapa contoh metode analisis yang diterapkan melalui jenis statistik ini antara lain:

a. Mengevaluasi signifikansi perbedaan upah minimum regional antara provinsi ataupun kabupaten / kota, dengan menerapkan analisis uji t.
b. Memahami pengaruh inflasi terhadap daya beli masyarakat, dengan menerapkan analisis regresi.
c. Mengidentifikasi hubungan antara tingkat suku bunga dengan jumlah peredaran uang, dengan menggunakan analisis korelasi.

### 1.2.4 Statistik non Parametrik

Melengkapi statistik parametrik, statistik non parametrik dibutuhkan jika jenis datanya menggunakan skala nominal atau skala ordinal. Termasuk juga jika kelompok data yang dianalisis tidak membutuhkan pemenuhan asumsi tertentu asumsi normalitas. Kelompok data dengan jumlah data yang tidak terlalu besar juga dapat menggunakan jenis statistik paramterik. Peran statistik non parametrik dalam

bidang ekonomi dan bisnis dapat berupa implementasi berikut:

a. Mengetahui tingkat kepuasan konsumen melalui kategori puas, netral dan tidak puas, dengan menggunakan analisis uji *Chi-Square.*
b. Menguji potensi perbedaan preferensi antara pria dan wanita dalam memilih karakteristik kendaraan, dengan menggunakan analisis uji *Mann-Whitney.*
c. Menguji signifikansi kesehatan perbankan melalui berbagai variabel keuangan pembanding, dengan menggunakan analisis uji Kruskal Wallis.

## 1.3 Jenis Data Statistik

Data merupakan landasan awal agar peran statistik dapat optimal sesuai dengan tujuan ataupun kebutuhan pengguna. Kualitas data sebelum sampai pada tahap analisis akan menentukan bagaimana cara data tersebut akan diuji dengan pendekatan yang sesuai. Data dalam statistik dapat diklasifikasikan berdasarkan beberapa kriteria, antara lain menurut sifatnya, sumber perolehannya dan waktu pengumpulannya.

### 1.3.1 Berdasarkan Sifat Data

1. Data Kualitatif

   Data kualitatif merupakan data yang tidak berbentuk angka dan umumnya jenis data ini menggambarkan karakteristik, deskripsi atau kategori tertentu. Data ini berupa kata-kata atau kalimat pendek sehingga tidak dapat diukur dengan tingkatan atau besaran nilai. Contoh data kualitatif antara lain jenis saham, persepsi konsumen, status sosial, status ekonomi, tingkat pendidikan.
2. Data Kuantitatif.

Data kuantitatif merupakan data yang dinyatakan dalam bentuk angka sehingga data ini dapat diukur besarannya secara numerik. Dengan kata lain, data ini dapat dikelola melalui perhitungan matematis. Data kuantitatif terbagi menjadi dua yaitu data diskrit dan data kontinu. Disebut data diskrit jika data tersebut berupa bilangan bulat atau tidak memiliki nilai pecahan, sebagai contoh, jumlah wilayah kabupaten/kota di Indonesia. Apabila data memiliki nilai desimal atau berada dalam rentang nilai tertentu disebut dengan data kontinu, sebagai contoh, persentase tingkat pertumbuhan ekonomi Indonesia.

### 1.3.2 Berdasarkan Sumber Perolehan

1. Data Primer

   Seluruh bentuk data yang dikumpulkan oleh peneliti atau enumerator (pengumpul data) secara langsung dari sumber utamanya didefenisikan sebagai data primer. Data primer sejatinya bersifat khusus atau spesifik sebab data yang dikumpulkan pasti disesuaikan dengan tujuan penelitian atau kebutuhan pengguna. Data primer dapat berasal dari beberapa cara, seperti:
   - Survei kepada nasabah terkait kepuasan layanan perbankan.
   - Wawancara langsung dengan pemilik atau pengelola UMKM (usaha mikro, kecil dan menengah) terkait strategi pemasaran digital yang diterapkan.
   - Observasi langsung terhadap perilaku konsumen pada bisnis ritel selama masa libur.
2. Data Sekunder

   Bentuk data lainnya yang dapat digunakan tetapi tidak dikumpulkan langsung dari sumber

utama maupun sumber pendukungnya didefenisikan sebagai data sekunder. Data tersebut telah ada sebelumnya, umumnya dalam bentuk laporan yang dipublikasikan, buku yang diterbitkan untuk publik, situs resmi dan dokumen legal pemerintah. Data sekunder juga dapat berasal dari beberapa hal, seperti:

- Publikasi atau berita resmi statistik oleh Badan Pusat Statistik.
- Jaringan Dokumentasi dan Informasi Hukum disingkat JDIH dari setiap lembaga negara.
- Lembaga publikasi ilmiah.

### 1.3.3 Berdasarkan Waktu Pengumpulan

1. *Data Cross Section*

   Data yang dikategorikan sebagai cross section apabila waktu pengumpulannya hanya berada pada satu ruang waktu tertentu (singkat) terhadap beberapa objek observasi. Objek observasi tidak hanya terbatas pada individu atau perusahaan tetapi dapat meluas menjadi area, daerah bahkan negara, tergantung pada kebutuhan pengguna. Salah satu contoh dari jenis data ini, data upah minimum regional dan pendapatan per kapita dari 10 kota besar di Indonesia pada tahun 2024.

2. *Data Time Series*

   Dikategorikan sebagai data time series, apabila data yang dikumpulkan berasal dari rentang waktu tertentu (panjang) terhadap satu objek obervasi. Data jenis ini sering dipergunakan untuk melakukan analisis peramalan atau meng-identifikasi pola pergerakan atau perubahan nilai data dari satu waktu ke waktu yg lain. Satu contoh dari data time series yaitu data fluktuasi harga

emas di Indonesia setiap triwulan mulai dari tahun 2015 hingga tahun 2024.

3. Data Panel

   Data panel merupakan kombinasi antara *data cross section* dengan data time series. Jenis data ini menyajikan kelompok data yang dikumpulan dari rentang waktu waktu tertentu terhadap beberapa objek observasi. Distribusi data ini juga dapat dikelola untuk menemukan *fixed effect* atau efek tetap dan *random effect* atau efek acak dari kelompok data yang dimiliki. Satu contoh dari data panel dapat berupa kelompok data pendapatan per kapita, tingkat pengangguran dan jumlah angkatan kerja dari 10 kota besar di Indonesia dalam rentang 10 tahun terakhir.

## 1.4 Skala Pengukuran Data

Skala pengukuran dapat dipahami sebagai pola atau rangkaian klasifikasi data yang disepakati untuk digunakan sebagai tolok ukur dalam menggambarkan hubungan dari nilai variabel yang menjadi objek observasi. Penentuan penggunaan jenis skala ukur yang sesuai dengan variabel akan menentukan teknik apa yang paling sesuai dalam tahapan pengumpulan data dan metode mana yang paling tepat digunakan dalam tahapan analisis data.

1. Skala Nominal

   Skala nominal merupakan skala pengukuran yang diperuntukkan hanya untuk mengelompokkan atau mengkategorikan tanpa memberlakukan tingkatan pada kelompok datanya. Data pada skala ini bersifat unik atau mutually exclusive, bermakna bahwa tidak terjadi nilai yang tumpang tindih antar kategori yang dibentuk. Skala nominal termasuk kedalam data kualitatif sehingga

tidak mengandung nilai numerik. Contoh dari skala nominal, mencakup:

- Jenis kelamin terbagi menjadi laki-laki dan perempuan
- Status perkawinan terbagi menjadi belum kawin, kawin, cerai
- Golongan darah terbagi menjadi A, B, AB, O

2. Skala Ordinal

Skala ordinal merupakan skala pengukuran yang diterapkan untuk mengelompokkan data kedalam peringkat atau tingkatan tertentu, akan tetapi jarak atau perbedaan dalam tingkatan tersebut belum memiliki nilai atau ukuran yang jelas. Ketidakpastian ukuran tersebut terjadi sebab selisih antar peringkat tidak memiliki proporsi yang sama besar. Skala ordinal juga termasuk kedalam data kualitatif sehingga tidak mengandung nilai numerik dan tidak memungkinkan dilakukan operasional secara matematis.

Contoh dari skala ordinal, antara lain:

- Tingkat pendidikan formal terbagi menjadi Sekolah Dasar, Sekolah Menengah Pertama, Sekolah Menengah Atas dan Pendidikan Tinggi.
- Peringkat dalam kompetisi terbagi menjadi juara 1, juara 2, juara 3 dan juara harapan.
- Variabel yang menggunakan likert scale atau skala likert. Skala likert terbagi menjadi dua klasifikasi yaitu favorable likert scale dan unfavorable likert scale. Kedua klasifikasi dipergunakan pada kondisinya masing-masing. Favorable berfokus pada pernyataan positif atau mendukung perspektif yang dibangun dalam variabel atau indikator sehingga pilihan jawaban dimulai dengan makna paling negatif dan diakhiri dengan makna paling positif. Sebaliknya, unfavorable berfokus

pernyataan negatif atau tidak mendukung perspektif yang dibangun, sehingga pilihan jawaban dimulai dengan makna paling positif dan diakhir dengan makna paling negatif. Jumlah pilihan jawaban dari skala likert juga beragam, paling sedikit 3 poin dan paling banyak 7 poin. Penggunaan pilihan skala ini misalnya untuk mengetahui tingkat kepuasan konsumen dengan 4 poin favorable likert scale, maka pilihan jawaban dapat terbagi menjadi sangat tidak puas, tidak puas, puas dan sangat puas.

Sebagai catatan, penggunaan skala ordinal seperti skala likert dalam mengukur tingkat persepsi yang akan dianalisis secara kuantitatif, tidak dapat dikuantifikasi secara langsung menjadi numerik, misalnya pada contoh diatas, sangat tidak puas bernilai 1 dan seterusnya hingga pilihan sangat puas bernilai 4. Hal ini akan menyebabkan bias dalam pengolahan data hingga hasil output analisis dan berpotensi menjadi interpretasi yang keliru. Oleh karena itu, transformasi data kualitatif menjadi kuantitatif dapat menggunakan bantuan *method of successive interval* (msi) sehingga perwakilan nilai numerik yang diberikan menjadi lebih proporsional.

3. Skala Interval

Skala interval merupakan skala ordinal yang telah memiliki peringkat atau tingkatan tertentu dengan jarak perbedaan atau selisih nilai yang sama besar dan konsisten. Akan tetapi, pada skala pengukuran ini tidak memiliki nilai nol absolut, dengan kata lain nilai nol tidak memiliki makna. Data dalam skala interval berbentuk angka dan termasuk kedalam data kuantitatif. Contoh dari skala interval, antara lain:

- Data untuk rentang waktu terbagi menjadi satuan jam, menit dan detik.

- *Semantic differensial scale* dapat digunakan dalam studi persepsi konsumen terhadap sebuah merk. Dengan menggunakan skala ini, responden akan disajikan pilihan berbentuk garis kontinum dengan dua pilihan pada kedua sisi paling ujung. Kedua pilihan tersebut berupa pasangan kata sifat bipolar yang saling bertolak belakang seperti "baik – buruk" atau "menyenangkan – membosankan". Panjang rentang yang diterapkan dalam skala ini umumnya adalah 5 poin, 7 poin atau 10 poin pilihan.
- *Numeric rating scale* dapat digunakan dalam studi persepsi konsumen terhadap inisiatif orientasi pelayanan. Skala ini menyajikan daftar numerik yang menjadi pilihan jawaban. Bersamaan dengan daftar tersebut, disampaikan deskripsi jelas sebagai makna dari setiap angka yang akan dipilih. Banyaknya daftar yang disampaikan, umumnya antara 5 hingga 10 pilihan, sesuai dengan tingkat kedalaman setiap indikator atau variabel.

4. Skala Rasio

Skala rasio merupakan skala interval yang telah memiliki nilai nol absolut, sehingga seluruh operasional matematis dimungkinkan untuk dilakukan. Skala ini juga wajib memiliki tingkatan atau peringkat dan nilai pembeda sama besar dan konsisten. Data dalam skala rasio juga berbentuk angka dan termasuk kedalam data kuantitatif.

Contoh dari skala rasion, antara lain:

- Data tingkat pendapatan atau besar laba perusahaan.
- Data tingkat kepadatan penduduk pada suatu area atau wilayah tertentu.
- Data pertumbuhan harga komoditas secara nasional.

## DAFTAR PUSTAKA

Agung, A. A. P., & Yuesti, A. (2019). Metode-Penelitian-Bisnis-Kuantitatif-Dan-Kualitatif. In *CV. Noah Aletheia* (Vol. 1, Issue 1).

Hamonangan, S., Karo, P. K., & Harahap, Z. (2021). Relationship Between Leadership Style with Service Quality. *Proceedings of the Palembang Tourism Forum 2021 (PTF 2021)*, *200*. https://doi.org/10.2991/aebmr.k.211223.015

Meilysa Pasaribu, R., Irvi Nurul Husna, A., Karo Karo, P., Hamonangan, S., Desy Rizkiyah, N., Ihdal Karomi, M., Subhan Iswahyudi, M., Elwadinata, F., Nurmadia Abdussamad, S., & Amar Jusman, I. (2025). *PENGARUH KUALITAS TERHADAP PELANGGAN - Google Buku*. Rey Media Grafika. https://books.google.co.id/books?hl=id&lr=&id=zjxHEQAAQBAJ&oi=fnd&pg=PA65&dq=related:R1bv4IBgZ8EJ:scholar.google.com/&ots=lNlNHyc5p-&sig=uVn1oZzEEvfkv-F9EffS7Px_L0k&redir_esc=y#v=onepage&q&f=false

Karo Karo, P., Hamonangan, S., & Zulkifli, A. A. (2023). *Pengolahan Data dengan SPSS* (1st ed., Vol. 1). Indomedia Pustaka. https://indomediapustaka.com/pengolahan-data-dengan-spss/

Karo Karo, P., Stiawan, M., & Darminiyanti. (2025). *PENGANTAR ILMU MANAJEMEN PEMASARAN - Google Books*. https://books.google.co.id/books?hl=en&lr=&id=55dNEQAAQBAJ&oi=fnd&pg=PA1&dq=info:4bPfoAtOxyoJ:scholar.google.com&ots=z9ZEndxlv4&sig=9fLPvOdhNV3j9_OGV6cMCEWF5is&redir_esc=y#v=onepage&q&f=false

Karo, P. K., Hamonangan, S., & Setiawan, A. (2022). The Analysis of Traditional Menu Promotion, Case Study: The City of Pagar Alam. *Pusaka: Journal of Tourism, Hospitality, Travel and Business Event*, *4*(2), 124–133. https://doi.org/10.33649/PUSAKA.V4I2.177

Sugiyono, P. D. (2018). Metode Penelitian Kuantitatif, Kualitatif, Kombinasi dan RD. *CV Alfabeta*.

# BAB 2

# PENGORGANISASIAN DAN PENYAJIAN DATA

**Oleh Anita Roosmawarni**

## 2.1 Pendahuluan

Volume data yang besar tentu memerlukan manajemen data yang baik. Manajemen data yang dimaksud mengacu pada proses pengorganisasian dan penyajian data dalam bentuk yang mudah dipahami untuk dianalisis dan digunakan. Sebagai contoh : (a) Mengapa Indonesia sulit sekali lepas dari jeratan impor bawang putih? Jawaban ini akan terjawab dari data Pusat Bioteknologi IPB yang menyatakan bahwa disparitas harga bawang putih impor dan lokal disebabkan karena biaya produksi bawang putih oleh petani lokal berada di level Rp. 30.000,-/kg. Hal ini pula yang mendorong harga bawang putih lokal menyentuh angka Rp. 43.700,-/kg sehingga kalah bersaing dengan bawang putih impor pada harga Rp. 18.400,-/kg; (b) Mengapa industri kreatif Indonesia perlu terus dikembangkan? Jawabannya mengacu pada data Dirjen Bea dan Cukai Kemenkeu, bahwa nilai ekspor ekonomi kreatif Indonesia pada tahun 2024 sebesar US $ 12,36 miliar. Artinya telah terjadi peningkatan 4.46% dari tahun sebelumnya sebagai dampak dari peningkatan permintaan ekspor kriya dan fashion di pasar global.

Contoh diatas merupakan bukti nyata bahwa data yang dikumpulkan akan bermanfaat apabila diorganisasikan dan disajikan dengan baik. Sehingga data mentah *(row data)* yang diperoleh dari populasi atau sampel menjadi data yang tertata rapi dan bermakna atau bernilai informasi bagi para pengambil keputusan managerial. Selain itu, data yang dikumpulkan dari lokasi penelitian pada umumnya belum teratur dan masih merupakan bahan keterangan yang sifatnya masih kasar dan sangat sedikit memberi informasi. Tugas statistik adalah mengorganisasikan dan menyajikannya dalam bentuk yang lebih ringkas dan mudah dimengerti (Pasaribu *et al.*, 2021).

Koleksi data statistika perlu disusun (diorganisir) dan disajikan (divisualkan) sedemikian rupa sehingga dapat dibaca dengan jelas dan mudah. Pada umumnya pengorganisasian dan penyajian data disajikan dalam bentuk : (a) tabel, (b) grafik, dan (c) diagram. Perlu diingat bahwasanya pengorganisasian dan penyajian data sangat bergantung kepada jenis data dan tujuan penyajiannya.

## 2.2 Tabel Data

Tabel merupakan sekumpulan angka/numerik yang disusun sedemikian rupa berdasarkan kategori tertentu. Tabel merupakan salah satu pilihan dalam pengorganisasian dan penyajian data yang sangat digemari oleh peneliti. Mengapa demikian? Karena dengan tabel, data dapat diorganisir dalam baris, kolom dan dilengkapi dengan identitas/keterangan sehingga pembaca akan mudah membandingkan dan mengidentifikasi pola data. Selain itu, data dalam bentuk tabel mempermudah proses pembacaan data, analisa data bahkan mempercepat proses membandingkan informasi.

### 2.2.1 Penyusunan Tabel Data

Data yang dikumpulkan dari lokasi penelitian harus ditata/diatur dengan baik agar memudahkan proses analisa. Berikut ini beberapa pilihan tahapan dalam penyusunan tabel data :

a. Teknik alfabetis

Teknik ini menyajikan data berdasarkan kolom nama yang diurutkan berdasarkan abjad/alfabet, dimulai dari abjad yang paling awal dalam kolom tersebut (*ascending*)

**Tabel 2. 1 Daftar Hadir Perkuliahan**

| NIM | Nama | Pert. 1 | Pert. 2 | Pert. 3 |
|---|---|---|---|---|
| 2025120 | Aan Saputra | √ | √ | |
| 2025121 | Clarissa Syakbani | √ | √ | √ |
| 2025123 | Dwy Evanda | | √ | |
| 2025125 | Hidayat Arifin | √ | | √ |

Sumber data : karangan

b. Teknik geografis

Teknik ini menampilkan data-data berdasarkan kolom nama yang diurutkan berdasarkan wilayah geografis.

**Tabel 2. 2 Pertumbuhan PDRB AHK Tahun 2020**

| Provinsi | PDRB |
|---|---|
| Aceh | 4.21 |
| Sumatera Utara | 4.73 |
| Sumatera Barat | 4.36 |

| Provinsi | PDRB |
|---|---|
| Riau | 4.55 |
| Jambi | 5.12 |
| Sumatera Selatan | 5.24 |
| Bengkulu | 4.31 |
| Lampung | 4.28 |

Sumber data : Badan Pusat Statistik, 2020

c. Teknik besaran angka

Teknik ini menyajikan tabel data yang angka-angkanya diurutkan mulai dari angka terkecil ke terbesar (*ascending*) maupun dari besar ke kecil (*descending*)

**Tabel 2. 3 Daftar Menu Makanan**

| No | Menu Makanan | Jumlah Pesanan |
|---|---|---|
| 1 | Lontong Sate | 1 |
| 2 | Soto Ayam | 10 |
| 3 | Lontong Balap | 5 |
| 4 | Nasi Rames | 7 |
| 5 | Iga Bakar | 15 |

Sumber data : karangan

d. Penyusunan secara historis

Tabel yang menyajikan data dengan mengklasifikasikan secara kronologis atau historis, biasanya dimulai dari waktu yang paling dahulu atau paling lama.

**Tabel 2. 4 Daftar Pertumbuhan Bunga Matahari**

| Minggu Ke- | Tanggal Pencatatan | Tinggi Tanaman |
|---|---|---|
| 1 | 1 Juni 2024 | 50 cm |
| 2 | 8 Juni 2024 | 75 cm |
| 3 | 13 Juni 2024 | 90 cm |
| 4 | 18 Juni 2024 | 120 cm |

e. Penyusunan atas klasifikasi tertentu

Tabel yang menyajikan data dengan menyusun pos-pos keterangan dalam tabel yang dilakukan atas dasar alasan statistik tertentu, misalkan tabel yang memuat informasi tentang aktivitas impor yang umumnya terbagi menjadi tiga kategori yakni (a) barang primer, (b) barang sekunder, dan (c) barang tersier.

### 2.2.2 Struktur Tabel

Struktur tabel yang baik harus bersifat sederhana dan jelas, agar memudahkan pembaca dalam memahami data yang disajikan dan tidak menimbulkan penafisran yang salah. Berikut adalah struktur tabel yang efisien :

1. Nama
   Tabel yang baik harus memiliki nama/titel sebagai identitas. Nama tabel harus jelas, singkat, dan padat dan pada umumnya mencirikan data yang disajikan. Nama tabel sebaiknya diletakkan di bagian atas tabel.
2. *Prefatory note* dan *foot note*
   *Prefatory* dan *foot note* tabel adalah bagian yang tidak dapat dipisahkan dari tabel. Keduanya

biasanya diposisikan dibawah nama tabel dan dibuat kurang menonjol dibandingkan nama tabel.

3. Sumber data
   Sumber data ditempatkan langsung dibawah tabel yang memuat informasi tentang asal-usul data didapatkan, baik data primer maupun sekunder.
4. Persentase
   Jika tabel menggunakan data-data dengan satuan persentase, maka setiap keterangan dalam tabel harus disajikan secara rinci dan jelas.
5. Jumlah
   Jika jumlah merupakan informasi yang dianggap penting, maka harus diletakkan pada posisi yang terlihat jelas, atau menuliskannya dalam huruf tebal. Apabila jumlah yang dimaksud memuat angka yang tidak terlalu penting maka dapat diletakkan di sisi kanan nama kolom.
6. Unit
   Beberapa data dalam tabel memerlukan informasi tentang unit pengukuran yang jelas, guna meminimalisasi salah penafsiran. Jika tidak maka ciri-ciri unit pengukurannya harus dijelaskan dalam nama kolom.

### 2.2.3 Jenis Tabel

Ketika proses pengorganisasian data dalam tabel telah dilakukan, maka secara umum tabel yang disajikan dapat diklasifikasikan dalam 4 bentuk tabel, yaitu :

1. Tabel Data Tidak Berkelompok
   Tabel data tidak berkelompok adalah tabel yang menyajikan data dalam bentuk titik-titik individual, seperti angka atau nilai. Tabel ini disebut juga tabel distribusi data tunggal.

**Tabel 2. 5 Tingkat Penghunian Kamar Pada Hotel Bintang Tahun 2024**

| Provinsi | Tingkat Hunian |
|---|---|
| Aceh | 41,58 |
| Sumatera Utara | 48,26 |
| Sumatera Barat | 44,34 |
| Riau | 45,09 |
| Dst... | |

Sumber : Badan Pusat Statistik Indonesia, 2024

2. Tabel Data Berkelompok
   Tabel data berkelompok adalah tabel yang menyajikan data dalam kelompok-kelompok atau kelas.

**Tabel 2. 6 Angka Harapan Lama Sekolah (HLS) Menurut Jenis Kelamin (Tahun), 2022-2024**

| Jenis Kelamin | 2022 | 2023 | 2024 |
|---|---|---|---|
| Laki-laki | 12.,96 | 12,98 | 12,99 |
| Perempuan | 13,28 | 13,33 | 13,39 |

Sumber data : Badan Pusat Statistik Indonesia, 2024

3. Tabel Data Kontingensi
   Tabel kontingensi merupakan bagian dari tabel data berkelompok, yang berisi data-data kategorikal yang membandingkan dua atau lebih variabel. Misalkan gunakan tabel kontingensi guna memahami hubungan antara jenis kelamin dengan jenis komputer yang digemari.

**Tabel 2. 7 Hubungan Antara Jenis Kelamin Dengan Jenis Komputer Yang Digemari**

| Jenis Kelamin | Jenis Komputer | | Jumlah |
|---|---|---|---|
| | PC | Mac | |
| Laki-laki | 66 | 40 | 106 |
| Perempuan | 30 | 87 | 117 |
| Jumlah | 96 | 127 | 223 |

Sumber data : data karangan

4. Tabel Data Frekuensi / Distribusi Frekuensi
   Tabel data frekuensi dilakukan dengan cara mengelompokan data ke dalam beberapa kategori. Tujuannya adalah menyederhanakan penyajian data sehingga lebih mudah dibaca dan dipahami. Perlu diingat bahwasanya setiap data hanya dapat dimasukkan ke dalam satu kategori. Adapun manfaat yang dapat diperoleh melalui metode ini adalah : (i) memudahkan dalam memahami pola distribusi data, (ii) memudahkan identifikasi nilai yang sering muncul, dan (iii) merangkum informasi statistik.

**Tabel 2. 8 Distribusi Pergerakan Saham di BEI Tahun 2023**

| Kelas ke- | Interval | Frekuensi | Jumlah Frekuensi |
|---|---|---|---|
| 1 | 77,0-84,4 | III | 3 |
| 2 | 84,5-91,9 | IIII | 4 |
| 3 | 92,0-99,4 | IIIIIII | 7 |
| 4 | 99,5-106,9 | IIII | 4 |
| 5 | 107,0-114,4 | IIII | 4 |

Sumber data : data karangan

## 2.3 Grafik / Diagram

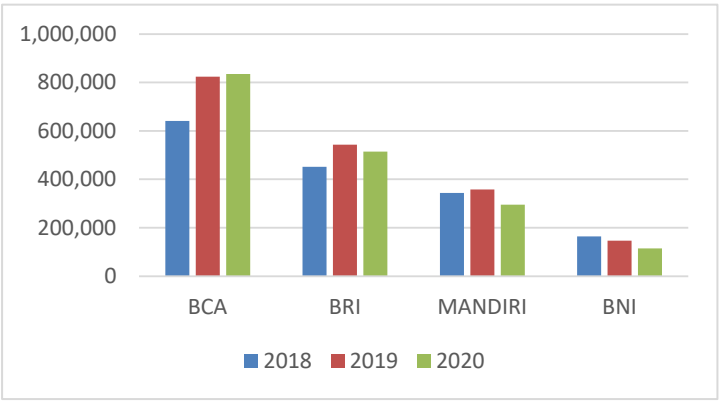
Grafiik / diagram merupakan penyajian informasi dalam bentuk visual. Metode ini digunakan dalam berbagai bidang seperti statistika, laporan keuangan perusahaan, riset-riset ilmiah, dll. Diagram merupakan salah satu metode penyajian data yang sangat disukai karena memeliki beberapa keunggulan yaitu :

a. Memudahkan pembacaan data
b. Visualisasi konsep yang kelas
c. Menarik perhatian pembaca
d. Meringkas informasi
e. Menyampaikan data, nilai, petunjuk atau prosedur
f. Mempermudah memahami konsep, ide, konstruksi, data statistik, anatomi data, dsb

### 2.3.1 Jenis Grafik / Diagram

Grafik / diagram yang digunakan dalam menyampaikan data dapat menggunakan beberapa jenis yang memiliki tampilan berbeda-beda. Berikut ini jenis-jenis grafik / diagram yang sering digunakan :

1. Grafik / diagram batang (histogram)
   Adalah adalah grafik yang tersusun dari segi empat yang menyerupai batang/balok kayu. Tinggi segi empat/batang menunjukkan jumlah frekuensi masing-masing kelas yang diwakili. (Pasaribu *et al.*, 2021).




**Gambar 2. 1 Data Perbankan Dengan Kapitalisasi Pasar Tertinggi Tahun 2018 – 2020**
(Sumber data : (Roosmawarni, Fatihudin and Mauliddah, 2022))

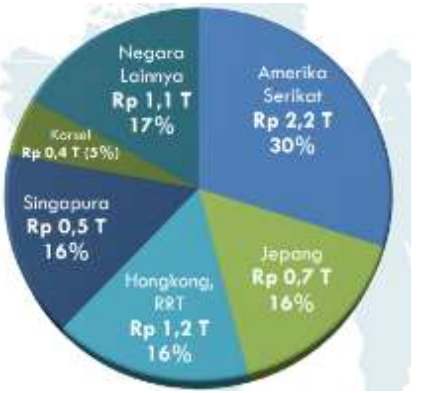
2. Diagram garis

Diagram garis (*line chart*) merupakan bentuk penyajian yang paling banyak dipakai dalam berbagai laporan baik perusahaan, pemerintahan, maupun penelitian ilmiah. Metode ini sering dipilih guna melihat trend/kecenderungan kenaikan/ penurunan atas data yang diolah. Salah satu bentuk data yang dapat diklasifikasi secara kronologis adalah data runtut waktu (*time series*) (Widyaningsih, 2021). Misalkan saja : sepanjang 2024, nilai tukar rupiah terhadap dolar Amerika Serikat (AS) mengalami fluktuasi (Indonesia, 2024)


**Gambar 2. 2 Pergerakan Nilai Rupiah Terhadap Dollar Sepanjang Tahun 2024**
(Sumber data : CNBC Indonesia, 2024)

3. Diagram lingkaran

Diagram lingkaran (*pie chart*) dapat dijadikan sebagai salah satu alternatif penyajian data dengan tampilan visual warna yang beragam. Diagram ini biasanya digunakan untuk menyatakan perbandingan jika data terdiri atas beberapa kelompok atau kategori. Sesuai dengan namanya, diagram ini berbentuk lingkaran yang dibagi dalam beberapa bagian dan memuat informasi di dalamnya. Pie chart dalam disajikan dalam bentuk 2 dimensi maupun 3 dimensi. Hal ini bisa membantu kita memvisualisasikan sebuah data agar dapat dengan mudah diterima oleh orang lain. Sebagai contoh, data tentang Realisasi PMA di Provinsi Jatim tahun 2022 yang didominasi oleh Amerika Serikat (sektor pertambangan) sebesar RP 5,5T, selanjutnya Hongkong dan RRC (pada industri logam, barag logam, bukan mesin dan peralatan sebesar Rp. 1,53T (BPS, 2022)




**Gambar 2. 3 Realisasi PMA Jawa Timur Triwulan 1 – 2022**

(Sumber data : BPS Provinsi Jatim, 2022)

4. Diagram Peta

   Diagram peta dapat dipilih sebagai salah satu metode penyajian data dalam bentuk gambar/wilayah tertentu yang melukiskan fenomena atau kejadian dihubungkan dengan suatu lokasi kejadian yang diamati (Widyaningsih, 2021). Misalkan diagram peta yang memvisualkan prakiraan puncak curah hujan pada tahun 2021-2022.

**Gambar 2. 4 Prakiraan Puncak Curah Hujan Provinsi Jatim**

(Sumber data : BPS Provinsi Jatim, 2022)

### 2.3.2 Struktur Grafik / Diagram

Struktur grafik / diagram yang benar, diupayakan memuat informasi berikut ini :

1. Jenis grafik/diagram yang dipilih

   Setiap jenis grafik/diagram memiliki keunggulan dan kekurangan masing-masing. Perlu diingat bahwa grafik yang baik tentunya memuat informasi yang mudah ditangkap oleh pembaca, bersifat sederhana, jelas, dan dapat ditarik kesimpulan. Grafik yang rumit biasanya disajikan untuk orang yang sangat mengerti permasalahan atau yang sangat mahir dalam ilmu statistik, dan informasi dalamgrafik tersebut tidak ditujukan untuk kepentingan publik.

2. Nama, skala sumbu, sumber, dan catatan lainnya

   Pencantuman nama, skala sumbu, sumber data, dan catatan lainnya wajib dilakukan dengan tujuan agar grafik/diagram yang disajikan sarat akan makna/informatif.

3. Pemberian tekanan pada grafik/diagram

   Penekanan tentang suatu peristiwa yang tertentu dalam penyajian grafik dapat dilakukan dengan cara memberi warna yang berbeda, tanda silang, atau garis yang berbeda. Garis dalam peta yang sama juga dapat dibedakan dengan menggunakan warna yang berbeda, garis terputus putus, garis padat (*solid line*) atau garis tebal. Garis padat lebih memberi 3tekanan dari pada garis terputus-putus, sedangkan garis tebal lebih menarik perhatian dari pada garis yang tipis.

## 2.4 Simpulan

Pengorganisasian dan penyajian data statistik dalam bentuk grafik/diagram, pada umumnya lebih menarik perhatian dan mengesankan dibandingkan tabel data. Grafik menyuguhkan visualisasi atas data yang diperoleh, dengan mengkolaborasi warna, jenis huruf, bentuk yang memanjakan mata. Inilah yang menyebabkan visual grafis lebih banyak digemari karena dapat melukiskan peristiwa secara mengesankan dan tidak membosankan.

Sebelum membuat visual grafis, analisis dan interpretasi data, sebaiknya data yang dikumpulkan telah tersusun/terformat dalam tabel statistik. Dengan kata lain, data yang berupa angka divisualkan. Berikut ini keunggulan penyajian data grafik :

a. Penyajian data grafis tampak lebih menarik karena mengkombinasikan warna, teks, dll
b. Grafik dapat dengan lebih cepat memperlihatkan gambaran umum secara menyeluruh tentang suatu perkembangan perubahan maupun perbandingan

c. Grafik yang dibuat dengan mengikuti kaidah yang benar, akan terasa lebih jelas dan lebih mudah dimengerti oleh pembaca.

Sebaliknya, grafis juga memiliki beberapa kelemahan dibandingkan tabel, antara lain :

a. Membuat grafik membutuhkan waktu pengerjaan yang lebih lama, biaya yang lebih mahal dan alat/instrumen yang lebih banyak
b. Data yang disajikan sangat terbatas. Jika data yang disajikan sangat banyak (*big data*), maka lukisan grafik menjadi tidak menarik, ruwet dan sangat membingungkan.
c. Ketelitian yang lebih rendah daripada tabel. Misalkan angka 10,00061; -19,32910 dapat dimuat dalam tabel namun tidak mungkin dilakukan pada grafik.

Dengan mengetahui berbagai keunggulan dan kelemahan tabel dan grafik sebagai alat pengorganisasian dan penyajian data, maka beberapa aspek diatas dapat dijadikan bahan pertimbangan dan pedoman untuk menetapkan apalah data yang sedang diteliti akan disajikan melalui tabel atau grafik.

## DAFTAR PUSTAKA

BPS (2022) *Data Dinamis Perekonomian Provinsi Jawa Timur Tahun 2021*. Available at: https://ro-ekonomi.jatimprov.go.id/storage/pdf/data dinamis/wDLdyB8r3ABmkuFoi3D4bF13VBiiqBmTsjg15kUy.pdf.

Indonesia, C. (2024) *Pergerakan Rupiah Sepanjang 2024, Bak Roller Coaster*, *CNBC Indonesia*. Available at: https://www.cnbcindonesia.com/research/20241229115444-128-599316/pergerakan-rupiah-sepanjang-2024-bak-roller-coaster.

Pasaribu, S. B. *et al.* (2021) *Statistika Ekonomi dan Bisnis*, *Metodologi Penelitian Untuk Ekonomi dan Bisnis*. Available at: https://library.idaqu.ac.id/index.php?p=show_detail&id=872&keywords=.

Roosmawarni, A., Fatihudin, D. and Mauliddah, N. (2022) 'Market Capitalisation and Financial Performance: Evidence from Banking Listed Company in Indonesia', *Jurnal Analisis Bisnis Ekonomi*, 20(2), pp. 124–136. doi: 10.31603/bisnisekonomi.v20i2.7835.

Widyaningsih, D. (2021) *Statistika Bisnis*. Edited by E. Zusrony. Semarang: Yayasan Prima Agusteknik.

# BAB 3
# UKURAN PEMUSATAN DATA

**Oleh Patrich Phill Edrich Papilaya**

## 3. 1 Pendahuluan

Dalam era data yang berkembang pesat saat ini, kemampuan untuk menginterpretasikan dan menganalisis informasi kuantitatif menjadi keterampilan krusial bagi para praktisi di bidang ekonomi dan bisnis (Grolemund & Wickham, 2017). Data yang melimpah, baik berskala besar maupun kecil, perlu diorganisasi dan diringkas secara sistematis untuk menghasilkan wawasan yang bermakna bagi pengambilan keputusan yang efektif (Davenport & Harris, 2017). Statistika, sebagai disiplin ilmu yang berfokus pada pengumpulan, analisis, dan interpretasi data, menawarkan berbagai metode untuk mengekstrak informasi penting dari kumpulan data yang kompleks (Montgomery & Runger, 2018). Salah satu konsep fundamental dalam statistika yang memiliki aplikasi luas dalam konteks ekonomi dan bisnis adalah ukuran pemusatan data.

Ukuran pemusatan data (*measures of central tendency*) merupakan nilai numerik yang menggambarkan kecenderungan pusat atau lokasi dari distribusi data (Weiss, 2019). Konsep ini berperan sebagai fondasi dalam statistika deskriptif yang bertujuan meringkas karakteristik utama dari sekumpulan data observasi (Healey, 2018). Dalam analisis ekonomi dan bisnis, ukuran pemusatan data membantu menyederhanakan informasi kompleks menjadi nilai tunggal

yang representatif, sehingga memfasilitasi perbandingan antar kelompok data dan pengambilan kesimpulan yang bermakna (Wooldridge, 2020).

Signifikansi ukuran pemusatan data dalam konteks ekonomi dan bisnis tidak dapat diremehkan. Dalam praktik manajemen, nilai-nilai ini menjadi dasar untuk benchmark kinerja, peramalan tren, evaluasi kebijakan, dan berbagai proses pengambilan keputusan strategis (Anderson et al., 2020). Sebagai contoh, rata-rata pendapatan per kapita sering digunakan sebagai indikator kesejahteraan ekonomi suatu negara, median harga properti dapat memberikan gambaran yang lebih representatif mengenai pasar real estate dibandingkan dengan mean yang rentan terhadap pengaruh nilai ekstrem, dan modus preferensi konsumen dapat menjadi informasi berharga dalam pengembangan produk dan strategi pemasaran (Berenson et al., 2018).

### 3.1.2 Evolusi Historis Konsep Pemusatan Data

Konsep pemusatan data memiliki sejarah panjang dalam perkembangan statistika sebagai disiplin ilmu. Gagasan tentang nilai tengah atau rata-rata dapat ditelusuri hingga ke peradaban kuno, di mana konsep-konsep primitif pemusatan data telah digunakan dalam aktivitas perdagangan, arsitektur, dan administrasi publik (Stigler, 2016). Astronom dan matematikawan Arab pada abad pertengahan, seperti Al-Biruni, telah mengembangkan metode penghitungan rata-rata untuk memperbaiki pengukuran astronomi (Porter, 2018). Perkembangan sistematis konsep pemusatan data dalam kerangka ilmiah modern mulai terjadi pada abad ke-17 dan ke-18 melalui kontribusi matematikawan seperti Blaise Pascal, Pierre de Fermat, dan Jakob Bernoulli dalam teori probabilitas (Stigler, 2016). Pada abad ke-19,

Adolphe Quetelet dan Francis Galton memperluas aplikasi konsep pemusatan data ke ranah ilmu sosial, sementara Karl Pearson dan Ronald Fisher pada awal abad ke-20 memformalisasi kerangka statistik modern yang menjadi fondasi analisis data saat ini (Porter, 2018).

Dalam konteks ekonomi dan bisnis, penggunaan ukuran pemusatan data menjadi semakin canggih sepanjang abad ke-20, dengan perkembangan teori ekonometrik oleh tokoh seperti Ragnar Frisch, Jan Tinbergen, dan Trygve Haavelmo (Wooldridge, 2020). Era komputasi dan revolusi digital telah lebih jauh memperluas kapabilitas analitik dan aplikasi ukuran pemusatan data dalam praktik ekonomi dan bisnis kontemporer (Davenport & Harris, 2017).

### 3.1.3 Relevansi dalam Konteks Ekonomi dan Bisnis Modern

Dalam lanskap ekonomi dan bisnis yang semakin kompleks dan data-*driven*, ukuran pemusatan data memainkan peran sentral dalam berbagai aplikasi (Provost & Fawcett, 2020). Beberapa area di mana konsep ini memiliki relevansi signifikan antara lain:

1) Analisis Pasar dan Industri: Ukuran pemusatan membantu mengidentifikasi tren sentral dalam variabel ekonomi seperti harga, volume perdagangan, dan nilai saham, memfasilitasi pemahaman dinamika pasar yang lebih baik (Brooks, 2019).
2) Manajemen Keuangan: Rata-rata tingkat pengembalian investasi (return on investment), median biaya operasional, dan modus perilaku pembayaran konsumen adalah beberapa contoh

aplikasi ukuran pemusatan dalam pengambilan keputusan keuangan (Berk & DeMarzo, 2021).

3) Ekonomi Makro: Indikator ekonomi makro seperti produk domestik bruto (PDB) per kapita, tingkat inflasi rata-rata, dan median pendapatan rumah tangga bergantung pada konsep pemusatan data untuk memberikan gambaran kondisi ekonomi secara keseluruhan (Mankiw, 2020).
4) Manajemen Sumber Daya Manusia: Struktur kompensasi, analisis produktivitas, dan penilaian kinerja karyawan sering menggunakan berbagai ukuran pemusatan sebagai benchmark dan alat evaluasi (Dessler, 2019).
5) Riset Pemasaran: Preferensi konsumen, pola belanja, dan respon terhadap strategi pemasaran dapat dianalisis menggunakan ukuran pemusatan untuk mengidentifikasi segmen pasar dan merancang kampanye yang efektif (Kotler & Keller, 2018).
6) Manajemen Operasi: Efisiensi proses produksi, pengendalian kualitas, dan optimasi rantai pasokan melibatkan analisis data berbasis ukuran pemusatan untuk mencapai performa optimal (Stevenson,2021).
7) Kebijakan Publik dan Ekonomi Pembangunan: Distribusi pendapatan, tingkat kemiskinan, dan efektivitas program sosial sering dianalisis menggunakan ukuran pemusatan untuk mengevaluasi dampak kebijakan ekonomi (Todaro & Smith, 2020).

Arti penting ukuran pemusatan data dalam konteks ini diilustrasikan dengan baik oleh Silver (2016) yang menekankan bahwa "kemampuan untuk meringkas dan menginterpretasikan distribusi data kompleks melalui ukuran pemusatan merupakan keterampilan

fundamental dalam ekonometrika modern dan analisis bisnis."

### 3.1.4 Jenis-jenis Ukuran Pemusatan Data

Terdapat beberapa ukuran pemusatan data yang umum digunakan dalam analisis statistik untuk konteks ekonomi dan bisnis, masing-masing dengan karakteristik, kelebihan, dan keterbatasan unik (Healey, 2018; Weiss, 2019):

1) Rata-rata Hitung (*Arithmetic Mean*) Rata-rata hitung atau yang sering disebut mean adalah ukuran pemusatan yang paling umum digunakan, dihitung dengan menjumlahkan semua nilai dalam dataset dan membaginya dengan jumlah observasi (Levine et al., 2017). Secara matematis, untuk set data $\{x_1, x_2, ..., x_n\}$, mean ($\mu$) didefinisikan sebagai: $\mu = (x_1 + x_2 + ... + x_n)/n = (\Sigma x_i)/n$ Mean memiliki sifat matematis yang baik dan memperhitungkan setiap nilai dalam dataset, menjadikannya ukuran yang komprehensif (Gravetter & Wallnau, 2019). Namun, mean sangat sensitif terhadap nilai ekstrem atau outlier, yang dapat mendistorsi representasi sentral dari data (Anderson et al., 2020). Dalam konteks ekonomi, mean sering digunakan untuk menganalisis variabel seperti pendapatan rata-rata, tingkat inflasi, dan return investasi (Wooldridge, 2020). Sebagai contoh, BPS (Badan Pusat Statistik) menggunakan pendapatan rata-rata untuk membantu menggambarkan kondisi ekonomi masyarakat Indonesia (BPS, 2021).
2) Median Median adalah nilai tengah dalam dataset yang telah diurutkan, membagi distribusi data menjadi dua bagian sama besar (Levine et al., 2017). Untuk dataset dengan jumlah observasi ganjil (n), median adalah nilai ke-((n+1)/2). Untuk

dataset dengan jumlah observasi genap, median adalah rata-rata dari dua nilai tengah. Keunggulan utama median adalah ketahanannya terhadap *outlier*, menjadikannya ukuran pemusatan yang lebih representatif untuk data yang tidak terdistribusi secara simetris (Healey, 2018). Karena karakteristik ini, median sering digunakan dalam analisis pendapatan, harga properti, dan variabel ekonomi lain yang menunjukkan skewness atau memiliki outlier (Berenson et al., 2018). Untuk ilustrasi, Bank Indonesia lebih memilih menggunakan median ekspektasi inflasi dalam survei ekonominya karena memberikan gambaran yang lebih akurat tentang ekspektasi umum dibandingkan dengan mean yang dapat dipengaruhi oleh respons ekstrem (Bank Indonesia, 2020).

3) Modus (*Mode*) Modus adalah nilai yang paling sering muncul dalam dataset (Weiss, 2019). Suatu distribusi dapat memiliki satu modus (unimodal), dua modus (bimodal), banyak modus (multimodal), atau tidak memiliki modus sama sekali. Modus sangat berguna untuk data kategorikal dan diskrit, serta dalam mengidentifikasi preferensi atau perilaku yang dominan (Gravetter & Wallnau, 2019). Dalam konteks bisnis, modus dapat menggambarkan produk terlaris, segmen pasar terbesar, atau waktu transaksi paling padat (Kotler & Keller, 2018). Sebagai contoh, perusahaan e-commerce seperti Tokopedia dan Shopee menganalisis modus perilaku belanja konsumen untuk mengoptimalkan strategi promosi dan penawaran produk (McKinsey, 2019).
4) Rata-rata Geometrik (*Geometric Mean*) Rata-rata geometrik dari n nilai positif adalah akar pangkat n dari hasil kali nilai-nilai tersebut (Berenson et al., 2018). Secara matematis: $G = \sqrt[n]{(x_1 \times x_2 \times \ldots \times x_n)} =$

$(\Pi x_i)^{(1/n)}$ Rata-rata geometrik sangat berguna untuk menganalisis tingkat pertumbuhan dan rasio dalam konteks ekonomi dan bisnis (Levine et al., 2017). Aplikasi umum termasuk menghitung tingkat pertumbuhan rata-rata (*compound annual growth rate* atau CAGR) untuk variabel ekonomi seperti GDP, investasi, atau nilai saham (Berk & DeMarzo, 2021). Otoritas Jasa Keuangan (OJK) dan lembaga investasi, misalnya, menggunakan rata-rata geometrik untuk melaporkan kinerja investasi jangka panjang, memberikan representasi yang lebih akurat dibandingkan rata-rata aritmetika (OJK, 2020).

5) Rata-rata Harmonik (*Harmonic Mean*) Rata-rata harmonik dari n nilai positif adalah n dibagi jumlah kebalikan dari nilai-nilai tersebut (Weiss, 2019). Secara matematis: $H = n / (1/x_1 + 1/x_2 + ... + 1/x_n) = n / \Sigma(1/x_i)$ Rata-rata harmonik berguna dalam analisis rasio dan tingkat, terutama ketika berhadapan dengan rata-rata dari rasio seperti kecepatan, harga per unit, atau efisiensi (Berenson et al., 2018). Dalam konteks bisnis, rata-rata harmonik dapat digunakan untuk menghitung rata-rata harga-pendapatan (*price-earnings ratio*) dari portofolio saham atau rata-rata tingkat pengembalian (Berk & DeMarzo, 2021).

### 3.1.5 Pemilihan Ukuran Pemusatan Data yang Tepat

Keputusan mengenai ukuran pemusatan mana yang paling sesuai untuk analisis tertentu bergantung pada berbagai faktor, termasuk jenis data, tujuan analisis, dan karakteristik distribusi (Healey, 2018). Beberapa pertimbangan utama dalam pemilihan ukuran pemusatan yang tepat meliputi:

1) Jenis Data: Data nominal lebih sesuai dianalisis dengan modus, data ordinal dengan median, sedangkan data interval dan rasio dapat menggunakan mean, median, atau modus tergantung pada distribusi dan konteks analisis (Weiss, 2019).
2) Distribusi Data: Untuk data yang terdistribusi normal atau simetris, mean biasanya menjadi pilihan yang tepat. Untuk data yang menceng (*skewed*) atau memiliki *outlier*, median sering memberikan representasi yang lebih baik (Anderson et al., 2020).
3) Tujuan Analisis: Tujuan dan konteks spesifik analisis juga memengaruhi pemilihan ukuran pemusatan. Sebagai contoh, analisis pendapatan sering menggunakan median untuk mengatasi ketidaksimetrisan, sementara analisis pengembalian investasi portofolio mungkin lebih tepat menggunakan rata-rata geometrik (Berk & DeMarzo, 2021).
4) Kehadiran *Outlier*: Keberadaan nilai ekstrem atau outlier dalam data seringkali mendorong penggunaan median daripada mean karena ketahanannya terhadap distorsi (Berenson et al., 2018).
5) Skala Pengukuran: Ukuran pemusatan yang berbeda memiliki properti matematis yang berbeda dan mungkin lebih sesuai untuk skala pengukuran tertentu (Gravetter & Wallnau, 2019).

### 3.1.6 Implikasi dalam Proses Pengambilan Keputusan

Pemahaman dan penerapan yang tepat dari ukuran pemusatan data memiliki implikasi signifikan dalam proses pengambilan keputusan ekonomi dan bisnis

(Davenport & Harris, 2017). Konsekuensi dari pemilihan ukuran pemusatan yang tidak tepat dapat meliputi:

1) Misrepresentasi Realitas: Penggunaan mean untuk data yang sangat menceng dapat memberikan gambaran yang menyesatkan tentang kecenderungan sentral yang sebenarnya (Silver, 2016).
2) Keputusan Suboptimal: Kebijakan ekonomi atau strategi bisnis yang didasarkan pada ukuran pemusatan yang tidak representatif dapat menghasilkan alokasi sumber daya yang tidak efisien dan keputusan suboptimal (Mankiw, 2020).
3) Bias dalam Evaluasi Kinerja: Evaluasi kinerja ekonomi atau bisnis menggunakan ukuran pemusatan yang tidak tepat dapat menghasilkan penilaian yang bias dan insentif yang tidak selaras (Dessler,2019).
4) Ekspektasi yang Tidak Realistis: Proyeksi dan peramalan berdasarkan ukuran pemusatan yang tidak tepat dapat menciptakan ekspektasi yang tidak realistis di kalangan pemangku kepentingan (Provost & Fawcett, 2020).

Untuk mengilustrasikan pentingnya pemilihan ukuran pemusatan yang tepat, dapat dipertimbangkan kasus klasik analisis pendapatan. Ketika mengevaluasi distribusi pendapatan dalam suatu populasi, penggunaan mean dapat menghasilkan estimasi yang terdistorsi ke atas karena pengaruh pendapatan sangat tinggi dari sebagian kecil individu. Median, dalam konteks ini, sering memberikan representasi yang lebih akurat tentang pendapatan "tipikal" (Todaro & Smith, 2020).

Sebagai contoh konkret, BPS Indonesia dalam Survei Sosial Ekonomi Nasional (SUSENAS) melaporkan baik mean maupun median pendapatan rumah tangga, dengan median secara konsisten lebih rendah dari mean, mencerminkan ketidaksimetrisan dalam distribusi pendapatan nasional (BPS, 2021). Perbedaan antara kedua ukuran ini sendiri menjadi indikator penting tentang tingkat ketimpangan ekonomi.

### 3.1.7 Tren Kontemporer dan Arah Masa Depan

Perkembangan teknologi big data, kecerdasan buatan, dan machine learning telah memperluas cakrawala aplikasi ukuran pemusatan data dalam analisis ekonomi dan bisnis (Provost & Fawcett, 2020). Beberapa tren kontemporer dan arah masa depan dalam domain ini meliputi:

1) Integrasi dengan Analitik Prediktif: Ukuran pemusatan semakin diintegrasikan dengan teknik analitik prediktif untuk menghasilkan wawasan yang lebih komprehensif dan prediksi yang lebih akurat (Shmueli et al., 2018).
2) Aplikasi dalam *Real-time Analytics*: Implementasi ukuran pemusatan dalam analitik real-time memungkinkan respons yang lebih cepat terhadap perubahan kondisi pasar dan perilaku konsumen (McAfee et al., 2020).
3) Perspektif Multidimensi: Pendekatan yang lebih holistik mengkombinasikan berbagai ukuran pemusatan dengan ukuran variabilitas dan bentuk distribusi untuk menangkap kompleksitas fenomena ekonomi dan bisnis secara lebih komprehensif (Hand, 2018).
4) *Customized Measures of Centrality*: Pengembangan ukuran pemusatan yang disesuaikan untuk konteks

dan dataset spesifik, melampaui ukuran konvensional, untuk menangkap aspek-aspek unik dari fenomena ekonomi dan bisnis (Mayer-Schönberger & Cukier, 2018).

Ukuran pemusatan data merupakan konsep fundamental dalam statistika yang memiliki aplikasi luas dan signifikan dalam analisis ekonomi dan bisnis. Melalui penyederhanaan distribusi data kompleks menjadi nilai representatif tunggal, ukuran ini memfasilitasi interpretasi, perbandingan, dan pengambilan keputusan berbasis data. Mean, median, modus, rata-rata geometrik, dan rata-rata harmonik masing- masing menawarkan perspektif unik dalam memahami kecenderungan sentral data, dengan kelebihan dan keterbatasan yang berbeda-beda.

Dalam lanskap ekonomi dan bisnis kontemporer yang semakin kompleks dan data-driven, pemahaman mendalam tentang berbagai ukuran pemusatan dan kemampuan untuk memilih ukuran yang paling sesuai untuk konteks analisis spesifik menjadi kompetensi yang semakin penting. Sebagaimana ditekankan oleh Davenport dan Harris (2017), "dalam era kompetisi berbasis analitik, keunggulan tidak hanya terletak pada ketersediaan data, tetapi pada kemampuan untuk mengekstrak wawasan bermakna melalui aplikasi metode statistik yang tepat."

Seiring dengan evolusi teknologi dan metode analitik, ukuran pemusatan data akan terus memainkan peran sentral dalam membantu praktisi ekonomi dan bisnis menavigasi kompleksitas fenomena kuantitatif, mengidentifikasi pola bermakna, dan membuat keputusan yang lebih informed dan efektif.

## 3.2 Mean (Rata-Rata)

### 3.2.1 Pengertian dan Konsep Dasar Mean

Mean atau rata-rata aritmatika merupakan ukuran pemusatan data yang paling umum digunakan dalam analisis statistik ekonomi dan bisnis. Secara konseptual, mean dapat didefinisikan sebagai nilai yang mewakili kecenderungan pusat dari sekumpulan data, yang diperoleh dengan menjumlahkan seluruh nilai dalam kumpulan data dan membaginya dengan banyaknya data (Lind et al., 2018). Mean memberikan informasi tentang nilai tipikal atau representatif dari suatu distribusi data.

Menurut Anderson et al. (2020), mean memiliki signifikansi khusus dalam analisis ekonomi dan bisnis karena kemampuannya dalam menyederhanakan interpretasi fenomena yang kompleks menjadi satu nilai tunggal yang memiliki makna statistik yang jelas. Mean juga memiliki kelebihan matematis karena melibatkan seluruh nilai dalam distribusi data dan menjadi dasar untuk analisis statistik lanjutan seperti analisis regresi, korelasi, dan inferensi statistik.

### 3.2.2 Jenis-Jenis Mean dan Formulasi Matematisnya

- **Mean Aritmatika (*Arithmetic Mean*)**

Mean aritmatika adalah jenis mean yang paling banyak digunakan dan biasanya menjadi rujukan utama ketika istilah "rata-rata" disebutkan tanpa kualifikasi lebih lanjut. Untuk data tidak berkelompok (*ungrouped data*), mean aritmatika dari sebuah sampel (dilambangkan dengan $\bar{x}$ dapat dihitung dengan rumus:

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

- $\bar{x}$ adalah mean aritmatika sampel
- $\sum_{i=1}^{n} x_i$ adalah jumlah seluruh nilai dalam kumpulan data
- n adalah banyaknya data

Sementara itu, untuk populasi rumusnya adalah:

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

Dimana:
- $\bar{x}$ adalah mean aritmatika populasi
- fi adalah jumlah seluruh nilai dalam populasi
- Xi adalah ukuran populasi

**Soal 1**

Sebuah perusahaan retail mencatat penjualan harian (dalam juta rupiah) dalam satu minggu sebagai berikut: 15, 12, 18, 20, 25, 30, 10. Berapakah rata-rata penjualan harian perusahaan tersebut?

**Jawaban:** Gunakan rumus rata-rata aritmatika:

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

$$\bar{X} = \frac{15 + 12 + 18 + 20 + 25 + 30 + 10}{7} = \frac{130}{7} = 18,57$$

Jadi, rata-rata penjualan harian perusahaan tersebut adalah 18,57 juta rupiah.

**Soal 2**

Seorang investor memiliki portofolio saham dengan komposisi: 40% saham A, 35% saham B, dan 25% saham C. Jika return saham A adalah 12%, saham B adalah 8%, dan saham C adalah 15%, berapakah return rata-rata tertimbang dari portofolio tersebut?

**Jawaban:** Gunakan rumus rata-rata tertimbang

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

Di mana:

- $w_i$ adalah bobot untuk nilai ke-i
- $X_i$ adalah nilai ke-i

$$\bar{X}_w = \frac{(0,40 \times 12\%) + (0,35 \times 8\%) + (0,25 \times 15\%)}{0,40 + 0,35 + 0,25}$$

$$\bar{X}_w = \frac{4,8\% + 2,8\% + 3,75\%}{1} = \frac{11,35\%}{1} = 11,35\%$$

$$\bar{X}_w = \frac{(50 \times 7) + (30 \times 15) + (15 \times 25) + (5 \times 50)}{50 + 30 + 15 + 5}$$

$$\bar{X}_w = \frac{350 + 450 + 375 + 250}{100} = \frac{1.425}{100} = 14,25$$

Jadi, rata-rata gaji keseluruhan karyawan di industri perbankan tersebut adalah 14,25 juta rupiah.

## 3.3 Median

Median adalah salah satu ukuran pemusatan data yang penting, terutama dalam analisis ekonomi dan bisnis di mana data seringkali tidak terdistribusi secara simetris. Berbeda dengan mean yang dihitung berdasarkan seluruh nilai data, median fokus pada nilai tengah dari kumpulan data yang telah diurutkan.

a. Sifat-sifat Median
- Tidak dipengaruhi oleh nilai ekstrem (*outlier*): Ini adalah keunggulan utama median dibandingkan mean. Jika ada nilai yang sangat tinggi atau sangat rendah dalam data, nilai median tidak akan banyak berubah.
- Unik: Setiap set data hanya memiliki satu nilai median.
- Dapat dihitung untuk data skala ordinal, interval, dan rasio. Meskipun paling sering digunakan untuk data interval dan rasio.
- Titik pusat secara posisi: Median adalah titik tengah secara fisik/posisi dalam data yang terurut.

b. Kelebihan dan Kekurangan Median

- Kelebihan:
  - Robust terhadap outlier: Sangat baik digunakan jika data memiliki nilai ekstrem.
  - Mudah diinterpretasikan: Konsep "nilai tengah" mudah dipahami.
  - Dapat dihitung untuk distribusi frekuensi dengan kelas terbuka (*open-ended classes*): Misalnya, jika kelas terakhir adalah "> 20 Juta Rupiah", median masih bisa dihitung selama kelas median tidak jatuh di kelas terbuka tersebut.
  - Cocok untuk data ordinal.
- Kekurangan:
  - Tidak menggunakan semua informasi data: Hanya nilai tengah (atau dua nilai tengah) yang digunakan secara langsung dalam perhitungan; nilai data lainnya hanya berfungsi untuk menentukan urutan.
  - Kurang sensitif: Perubahan pada nilai data yang bukan di tengah (selama urutan tidak berubah) tidak akan mengubah median.
  - Lebih sulit diolah secara matematis: Dibandingkan mean, median kurang cocok untuk analisis statistik inferensial yang lebih lanjut (misalnya, dalam pengujian hipotesis tertentu).
  - Perhitungan untuk data berkelompok bisa lebih rumit daripada menghitung mean.

Median menjadi pilihan yang lebih baik daripada mean dalam situasi berikut:

- ✓ Distribusi Data Miring (*Skewed*): Untuk distribusi pendapatan, harga rumah, atau waktu respons, di mana beberapa nilai tinggi dapat "menarik" mean ke arah ekor distribusi, median memberikan gambaran pusat yang lebih representatif.
- ✓ Adanya Nilai Ekstrem (*Outliers*): Jika ada pencilan data yang signifikan, mean akan sangat terpengaruh, sedangkan median tetap stabil.
- ✓ Data Ordinal: Ketika data hanya menunjukkan urutan (misal: peringkat kepuasan: sangat tidak puas, tidak puas, netral, puas, sangat puas), median adalah ukuran pemusatan yang paling tepat.
- ✓ Distribusi Frekuensi dengan Kelas Terbuka: Jika batas atas kelas terakhir atau batas bawah kelas pertama tidak ditentukan.

### 3.3.1 Definisi dan Cara Menentukan Median

Median (Me) adalah nilai yang membagi suatu set data yang telah diurutkan (dari terkecil ke terbesar atau sebaliknya) menjadi dua bagian yang sama besar. Artinya, 50% data akan berada di bawah nilai median, dan 50% data akan berada di atas nilai median. Cara menentukan median bergantung pada apakah data tersebut merupakan data tunggal (*ungrouped data*) atau data berkelompok (*grouped data*).

- ➢ Data Tunggal (*Ungrouped Data*)
  1. Urutkan Data: Langkah pertama dan paling krusial adalah mengurutkan data dari nilai terkecil hingga terbesar.

2. Tentukan Posisi Median:

- Jika jumlah data (n) ganjil: Median adalah nilai data yang berada tepat di tengah, yaitu pada posisi ke-(n + 1) / 2.
- *Contoh:* Data penjualan harian (dalam unit): 15, 20, 12, 25, 18.
  - Urutkan: 12, 15, **18**, 20, 25 (n=5)
  - Posisi Median = (5 + 1) / 2 = 3.
  - Median (Me) adalah data ke-3, yaitu **18 unit**.
- Jika jumlah data (n) genap: Median adalah rata-rata dari dua nilai data yang berada di tengah, yaitu data pada posisi ke-(n / 2) dan data pada posisi ke-((n / 2) + 1).
- *Contoh:* Data skor kepuasan pelanggan (skala 1-10): 7, 8, 5, 9, 6, 10.
  - Urutkan: 5, 6, **7**, **8**, 9, 10 (n=6)
  - Posisi tengah adalah data ke-(6/2) = 3 dan data ke-((6/2)+1) = 4.
  - Median (Me) adalah rata-rata dari data ke-3 (nilai 7) dan data ke-4 (nilai 8).
  - Me = (7 + 8) / 2 = **7.5**.

➢ Data Berkelompok (*Grouped Data*)

Ketika data disajikan dalam bentuk tabel distribusi frekuensi, median diestimasi menggunakan formula berikut:

**Me = L + [ ((n / 2) - F) / f ] * c**

Di mana:

- **Me** = Median
- **L** = Tepi bawah kelas median (Lower boundary of the median class). Kelas median adalah kelas interval di mana nilai median terletak (kelas di

mana frekuensi kumulatif mencapai atau melewati n/2).

- **n** = Jumlah total frekuensi (total data).
- **F** = Frekuensi kumulatif *sebelum* kelas median (*Cumulative frequency before the median class*).
- **f** = Frekuensi kelas median (*Frequency of the median class*).
- **c** = Panjang interval kelas (*Class interval width*).

Langkah-langkah Menentukan Median Data Berkelompok:

1. Hitung total frekuensi (n).
2. Tentukan posisi median: n / 2.
3. Buat kolom frekuensi kumulatif (kurang dari).
4. Identifikasi kelas median: kelas pertama di mana frekuensi kumulatifnya sama dengan atau lebih besar dari n / 2.
5. Tentukan nilai L, F, f, dan c dari kelas median tersebut.
6. Masukkan nilai-nilai ke dalam formula median.

**Soal**

Distribusi Pendapatan Bulanan Karyawan (dalam Juta Rupiah)

| **Pendapatan (Juta Rp)** | **Frekuensi (f)** | **Frekuensi Kumulatif (F Kum)** |
|---|---|---|
| 5 - < 7 | 8 | 8 |
| 7 - < 9 | 12 | 20 |
| 9 - < 11 | 15 | 35 <-- **Kelas Median** |
| 11 - < 13 | 10 | 45 |
| 13 - < 15 | 5 | 50 |
| **Total** | **n = 50** | |

1. n = 50
2. Posisi median = n / 2 = 50 / 2 = 25. Data ke-25 terletak di suatu tempat dalam distribusi.
3. Frekuensi kumulatif dibuat (lihat tabel).
4. Kelas median: Frekuensi kumulatif mencapai 25 (atau lebih) pertama kali pada kelas 9 - < 11 (karena F Kum = 35 > 25, dan F Kum sebelumnya adalah 20 < 25).
5. Identifikasi nilai:
   - L = Tepi bawah kelas median = 9
   - n = 50
   - F = Frekuensi kumulatif sebelum kelas median = 20
   - f = Frekuensi kelas median = 15
   - c = Panjang interval kelas = 11 - 9 = 2

6. Hitung Median:
   Me = 9 + [ ( (50 / 2) - 20 ) / 15 ] * 2
   Me = 9 + [ ( 25 - 20 ) / 15 ] * 2
   Me = 9 + [ 5 / 15 ] * 2
   Me = 9 + (1/3) * 2
   Me = 9 + 0.67
   Me ≈ 9.67 (Juta Rupiah)

   *Interpretasi*: Diperkirakan 50% karyawan memiliki pendapatan bulanan kurang dari 9.67 Juta Rupiah, dan 50% lainnya memiliki pendapatan lebih dari 9.67 Juta Rupiah.

## 3.4 Modus (Mode)

Selain mean dan median, modus adalah ukuran pemusatan data lainnya yang memberikan perspektif berbeda tentang 'pusat' atau 'kecenderungan umum' dari suatu kumpulan data. Modus secara khusus menyoroti nilai atau kategori yang paling sering muncul dalam suatu dataset (Lind et al., 2021).

a. Sifat-sifat Modus

   Modus memiliki beberapa sifat kunci (Levine et al., 2019; Lind et al., 2021):

   - Tidak selalu unik; bisa jadi bimodal atau multimodal.
   - Mungkin tidak ada sama sekali.
   - Tidak dipengaruhi oleh nilai ekstrem (outlier), sama seperti median.
   - Merupakan satu-satunya ukuran pemusatan yang cocok untuk data berskala nominal.
   - Menunjukkan nilai yang paling 'populer' atau 'tipikal' berdasarkan frekuensi.

b. Kelebihan dan Kekurangan Modus

- Kelebihan:
  - Mudah dipahami dan ditentukan (terutama untuk data tunggal) (Levine et al., 2019).
  - Merepresentasikan nilai yang paling sering muncul atau paling populer.
  - Tidak terpengaruh oleh nilai ekstrem (Lind et al., 2021).
  - Dapat digunakan untuk semua skala pengukuran data, termasuk nominal (Keller, 2017).
- Kekurangan:
  - Tidak selalu ada atau tidak selalu unik (Levine et al., 2019).
  - Tidak mempertimbangkan semua nilai dalam data.
  - Bisa jadi tidak stabil; perubahan kecil pada data dapat mengubah modus secara drastis.
  - Kurang berguna untuk analisis statistik inferensial lebih lanjut (Lind et al., 2021).
  - Estimasi modus untuk data berkelompok hanyalah perkiraan (Keller, 2017).

c. Kapan Modus Digunakan?

Modus paling berguna dalam situasi berikut (Keller, 2017; Levine et al., 2019):

- Ketika bekerja dengan data kualitatif/kategorikal, terutama data nominal, untuk mengetahui kategori yang paling umum.
- Untuk mengidentifikasi puncak distribusi atau nilai yang paling sering terjadi dalam data kuantitatif.
- Sebagai deskripsi cepat tentang nilai yang paling umum.
- Dalam manufaktur dan kontrol kualitas untuk menentukan spesifikasi produk yang paling umum.

### 3.4.1 Definisi dan Cara Menentukan Modus

Modus (Mo) adalah nilai data yang memiliki frekuensi kemunculan tertinggi dalam suatu kumpulan data (Keller, 2017). Dengan kata lain, modus adalah nilai atau kategori yang paling sering terjadi. Cara menentukan modus berbeda untuk data tunggal dan data berkelompok.

- Data Tunggal (*Ungrouped Data*)
  Menentukan modus untuk data tunggal relatif mudah: Cukup identifikasi nilai data yang memiliki frekuensi tertinggi setelah menghitung kemunculan setiap nilai (Levine et al., 2019).
  - *Contoh 1 (Unimodal):* Data jumlah cacat per batch produksi: 2, 3, 1, 3, 4, 3, 2, 5.
    - Frekuensi: 1 (muncul 1x), 2 (muncul 2x), 3 (muncul 3x), 4 (muncul 1x), 5 (muncul 1x).
    - Modus (Mo) = 3, karena nilai 3 muncul paling sering.
  - *Contoh 2 (Bimodal):* Data preferensi rasa minuman (A, B, C): A, B, B, C, A, B, C, A, B, A.
    - Frekuensi: A (muncul 4x), B (muncul 4x), C (muncul 2x).
    - Data ini memiliki dua modus (bimodal): A dan B.
  - *Contoh 3 (No Mode):* Data rating layanan (skala 1-5): 1, 2, 3, 4, 5.
    - Setiap nilai muncul hanya satu kali. Data ini tidak memiliki modus.
- Data Berkelompok (*Grouped Data*)

Untuk data yang disajikan dalam tabel distribusi frekuensi, modus tidak dapat ditentukan secara pasti, tetapi dapat diestimasi dengan mengidentifikasi kelas modus (kelas dengan frekuensi tertinggi) terlebih dahulu (Lind et al., 2021). Estimasi nilai modus di dalam kelas tersebut menggunakan formula berikut:

$$\mathbf{Mo = L + [\ d1\ /\ (d1 + d2)\ ]\ *\ c}$$

(Lind et al., 2021; Levine et al., 2019)
Di mana:

- **Mo** = Modus (estimasi)
- **L** = Tepi bawah kelas modus (*Lower boundary of the modal class*).
- **d1** = Selisih antara frekuensi kelas modus dengan frekuensi kelas *sebelumnya*.
- **d2** = Selisih antara frekuensi kelas modus dengan frekuensi kelas *sesudahnya*.
- **c** = Panjang interval kelas (*Class interval width*).

**Contoh:**

Distribusi Waktu Tunggu Pelanggan di Bank (dalam Menit)

| Waktu Tunggu (Menit) | Frekuensi (f) |
|---|---|
| 0 - < 5 | 10 |
| 5 - < 10 | 18 |
| 10 - < 15 | **25** |
| 15 - < 20 | 15 |
| 20 - < 25 | 7 |

| | |
|---|---|
| **Total** | **75** |

1. Kelas modus: 10 - < 15 (frekuensi tertinggi = 25).
2. Identifikasi nilai: L=10, f_modus=25, f_sebelum=18, f_sesudah=15, d1=25-18=7, d2=25-15=10, c=5.
3. Hitung Modus:
   Mo = 10 + [ 7 / (7 + 10) ] * 5
   Mo = 10 + [ 7 / 17 ] * 5
   Mo = 10 + (0.4118) * 5
   Mo = 10 + 2.059
   Mo ≈ 12.06 (Menit)

***Interpretasi**:* Berdasarkan data kelompok ini, diperkirakan waktu tunggu pelanggan yang paling sering terjadi adalah sekitar 12.06 menit

## 3.5 Perbandingan Mean, Median, dan Modus

Setelah memahami masing-masing ukuran pemusatan data (mean, median, dan modus), penting untuk mengetahui bagaimana ketiganya berhubungan satu sama lain dan kapan sebaiknya menggunakan masing-masing ukuran tersebut. Pilihan ukuran pemusatan yang tepat sangat bergantung pada karakteristik data, terutama bentuk distribusinya dan adanya nilai ekstrem (*outlier*).

### 3.5.1 Hubungan antara Mean, Median, dan Modus pada Berbagai Bentuk Distribusi

Bentuk distribusi frekuensi data memberikan petunjuk visual mengenai hubungan antara mean, median, dan modus (Lind et al., 2021).

- Distribusi Simetris (*Symmetrical Distribution*)
  - Deskripsi: Distribusi simetris memiliki bentuk yang seimbang di kedua sisi titik tengahnya. Jika dilipat pada titik tengah, kedua sisinya akan saling menutupi. Contoh paling umum adalah distribusi normal (kurva lonceng).
  - Hubungan: Pada distribusi yang *benar-benar* simetris, ketiga ukuran pemusatan data akan bernilai sama atau sangat berdekatan: Mean = Median = Modus. Ketiganya terletak di pusat distribusi, yang juga merupakan titik puncak kurva (Keller, 2017).
  - *[Ilustrasi Konseptual: Bayangkan kurva lonceng yang sempurna, puncaknya adalah modus, titik tengah area di bawah kurva adalah median, dan pusat keseimbangan data adalah mean. Ketiganya berada di titik yang sama.]*
- Distribusi Miring ke Kanan (*Positively Skewed Distribution*)
  - Deskripsi: Distribusi ini memiliki 'ekor' yang memanjang ke arah kanan (nilai-nilai yang lebih besar). Puncak distribusi berada di sisi kiri, dan sebagian besar data terkonsentrasi di nilai-nilai yang lebih rendah, namun ada beberapa nilai ekstrem yang tinggi. Contoh umum adalah data pendapatan, harga rumah, atau waktu tunggu.
  - Hubungan: Pada distribusi miring ke kanan, urutan ukuran pemusatannya adalah: Modus < Median < Mean.
    - Modus: Terletak pada puncak kurva (nilai paling sering muncul), yang berada di sisi kiri.

- Median: Terletak di tengah data *setelah diurutkan*, sehingga posisinya sedikit tertarik ke kanan (ke arah ekor) dibandingkan modus.
- Mean: Sangat dipengaruhi oleh nilai-nilai ekstrem yang tinggi di ekor kanan, sehingga nilainya paling 'tertarik' ke kanan dan menjadi yang terbesar di antara ketiganya (Levine et al., 2019; Lind et al., 2021).

  - *[Ilustrasi Konseptual: Bayangkan kurva dengan puncak di kiri dan ekor panjang ke kanan. Puncak adalah Modus. Median membagi area menjadi dua, sedikit ke kanan dari puncak. Mean ditarik lebih jauh ke kanan oleh nilai-nilai tinggi di ekor.]*

- Distribusi Miring ke Kiri (*Negatively Skewed Distribution*)
  - Deskripsi: Distribusi ini memiliki 'ekor' yang memanjang ke arah kiri (nilai-nilai yang lebih kecil). Puncak distribusi berada di sisi kanan, dan sebagian besar data terkonsentrasi di nilai-nilai yang lebih tinggi, namun ada beberapa nilai ekstrem yang rendah. Contohnya bisa berupa skor ujian yang relatif mudah (banyak nilai tinggi).
  - Hubungan: Pada distribusi miring ke kiri, urutan ukuran pemusatannya adalah kebalikan dari distribusi miring ke kanan: Mean < Median < Modus.
    - Modus: Terletak pada puncak kurva (nilai paling sering muncul), yang berada di sisi kanan.

- Median: Terletak di tengah data *setelah diurutkan*, posisinya sedikit tertarik ke kiri (ke arah ekor) dibandingkan modus.
- Mean: Sangat dipengaruhi oleh nilai-nilai ekstrem yang rendah di ekor kiri, sehingga nilainya paling 'tertarik' ke kiri dan menjadi yang terkecil di antara ketiganya (Levine et al., 2019; Keller, 2017).

- *[Ilustrasi Konseptual: Bayangkan kurva dengan puncak di kanan dan ekor panjang ke kiri. Puncak adalah Modus. Median membagi area menjadi dua, sedikit ke kiri dari puncak. Mean ditarik lebih jauh ke kiri oleh nilai-nilai rendah di ekor.]*

### 3.5.2 Panduan Memilih Ukuran Pemusatan Data yang Tepat

Memilih ukuran pemusatan yang paling sesuai adalah kunci untuk interpretasi data yang akurat. Berikut panduannya (Keller, 2017; Lind et al., 2021):

| Kriteria Pertimbangan | Mean (Rata-rata) | Median | Modus |
|---|---|---|---|
| **Skala Pengukuran Data** | Interval, Rasio | Ordinal, Interval, Rasio | Nominal, Ordinal, Interval, Rasio |
| **Bentuk Distribusi** | **Simetris:** Pilihan baik, representatif. | **Simetris:** Representatif. **Miring (Skewed** | **Simetris/Miring:** Kurang representatif dibanding Mean/Media |

| Kriteria Pertimbangan | Mean (Rata-rata) | Median | Modus |
|---|---|---|---|
| | | **): Pilihan terbaik.** | n. |
| **Keberadaan Outlier** | **Sangat terpengaruh.** Hindari jika ada outlier signifikan. | **Tidak terpengaruh (Robust). Pilihan terbaik** jika ada outlier. | **Tidak terpengaruh.** Bisa jadi pilihan, tapi kurang informatif untuk data numerik. |
| **Tujuan Analisis** | Deskripsi pusat data (jika simetris), dasar untuk analisis statistik lanjutan (inferensial). | Deskripsi titik tengah data, representasi 'tipikal' pada data miring/outlier. | Identifikasi nilai/kategori paling umum/populer. Cocok untuk data nominal. |
| **Karakteristik Lain** | Memanfaatkan semua nilai data. | Hanya posisi tengah yang penting. Bisa untuk kelas terbuka. | Mungkin tidak unik atau tidak ada. Paling mudah untuk data kategorikal. |

- Ringkasan Keputusan:
  - Gunakan Mean jika data berskala interval/ rasio, distribusinya mendekati simetris, dan tidak ada outlier signifikan. Mean baik untuk analisis statistik lebih lanjut.
  - Gunakan Median jika data berskala ordinal, interval, atau rasio; terutama jika distribusinya miring (*skewed*) atau terdapat outlier signifikan. Median memberikan gambaran nilai tengah yang lebih stabil dalam kondisi ini.
  - Gunakan Modus terutama untuk data berskala nominal, atau ketika tujuan utamanya adalah mengetahui nilai atau kategori yang paling sering muncul.

### 3.5.3 Ilustrasi Perbandingan Menggunakan Studi Kasus Ekonomi/Bisnis

Mari kita lihat contoh data gaji bulanan (dalam Juta Rupiah) 7 orang karyawan di sebuah divisi kecil:
Data Gaji: 8, 9, 8, 10, 11, 9, 35

1. Urutkan Data: 8, 8, 9, 9, 10, 11, 35
2. Hitung Mean:
   Mean = (8 + 8 + 9 + 9 + 10 + 11 + 35) / 7
   Mean = 90 / 7
   Mean ≈ 12.86 (Juta Rupiah)
3. Hitung Median:
   Jumlah data (n) = 7 (ganjil). Posisi median = (7 + 1) / 2 = 4.
   Data ke-4 adalah 9.
   Median = 9 (Juta Rupiah)

4. Hitung Modus:
   Nilai 8 muncul 2 kali. Nilai 9 muncul 2 kali. Nilai lainnya muncul 1 kali.
   Data ini memiliki dua modus (bimodal).
   Modus = 8 dan 9 (Juta Rupiah)

- Analisis:
  - Dalam kasus ini, Mean (12.86 Juta) jauh lebih tinggi daripada Median (9 Juta) dan Modus (8 & 9 Juta).
  - Penyebabnya adalah adanya satu nilai gaji yang sangat tinggi (outlier), yaitu 35 Juta. Nilai ekstrem ini 'menarik' nilai mean ke atas secara signifikan.
  - Median (9 Juta) tampaknya menjadi representasi yang lebih baik untuk gaji 'tipikal' karyawan di divisi tersebut, karena nilai ini berada di tengah-tengah setelah data diurutkan dan tidak terpengaruh oleh gaji yang sangat tinggi.
  - Modus (8 dan 9 Juta) menunjukkan bahwa gaji 8 Juta dan 9 Juta adalah yang paling sering muncul, memberikan informasi tentang frekuensi tetapi mungkin kurang mewakili pusat data secara keseluruhan dibandingkan median dalam kasus ini.

- Kesimpulan Ilustrasi: Jika manajer melaporkan bahwa rata-rata (mean) gaji di divisi tersebut adalah 12.86 Juta, ini bisa memberikan gambaran yang menyesatkan tentang kondisi sebagian besar karyawan. Melaporkan median (9 Juta) akan memberikan pemahaman yang lebih akurat mengenai tingkat gaji yang umum diterima oleh karyawan di divisi tersebut. Ini menunjukkan pentingnya memilih ukuran pemusatan yang tepat sesuai konteks dan karakteristik data.

## 3.6 Ukuran Letak (Kuartil, Desil, Persentil)

Setelah membahas ukuran pemusatan data (mean, median, modus) yang menggambarkan pusat distribusi, kita beralih ke Ukuran Letak atau Ukuran Posisi. Ukuran-ukuran ini, seperti kuartil, desil, dan persentil, memberikan informasi tentang posisi relatif suatu nilai dalam keseluruhan kumpulan data yang telah diurutkan. Mereka merupakan pengembangan dari konsep median dan membantu memahami penyebaran serta peringkat data secara lebih rinci (Lind et al., 2021).

### 3.6.1 Kuartil (*Quartiles*)

- Definisi

  Kuartil adalah nilai-nilai yang membagi sekumpulan data yang telah diurutkan menjadi empat bagian yang sama besar. Ada tiga titik kuartil:
  - Kuartil Pertama (Q1) atau Kuartil Bawah: Nilai yang membatasi 25% data terendah dari 75% data tertinggi. Artinya, 25% data berada di bawah Q1. Q1 setara dengan Persentil ke-25 (P25).
  - Kuartil Kedua (Q2): Nilai yang membagi data menjadi dua bagian sama besar (50% di bawah, 50% di atas). Q2 sama dengan Median. Q2 setara dengan Persentil ke-50 (P50) dan Desil ke-5 (D5).

- Kuartil Ketiga (Q3) atau Kuartil Atas: Nilai yang membatasi 75% data terendah dari 25% data tertinggi. Artinya, 75% data berada di bawah Q3 (atau 25% data berada di atas Q3). Q3 setara dengan Persentil ke-75 (P75).

➢ Perhitungan Kuartil Data Tunggal (*Ungrouped Data*)

1. Urutkan Data: Susun data dari nilai terkecil hingga terbesar.
2. Tentukan Posisi Kuartil (Qi): Gunakan formula Posisi(Qi) = i * (n + 1) / 4, di mana 'i' adalah kuartil ke- (1, 2, atau 3) dan 'n' adalah jumlah data.
3. Tentukan Nilai Kuartil:
   - Jika posisi Qi adalah bilangan bulat, nilai Qi adalah data pada posisi tersebut.
   - Jika posisi Qi adalah bilangan pecahan (misal, k.d), nilai Qi diestimasi dengan interpolasi: Nilai Qi = Data ke-k + d * (Data ke-(k+1) - Data ke-k), di mana 'k' adalah bagian bulat dan 'd' adalah bagian desimal dari posisi Qi.

- *Contoh:* Data waktu penyelesaian tugas (menit): 12, 15, 18, 20, 22, 25, 28, 30, 35, 40, 42 (n=11).
  - Data sudah terurut.
  - Posisi Q1: 1 * (11 + 1) / 4 = 12 / 4 = 3. Nilai Q1 adalah data ke-3, yaitu 18.
  - Posisi Q2 (Median): 2 * (11 + 1) / 4 = 24 / 4 = 6. Nilai Q2 adalah data ke-6, yaitu 25.
  - Posisi Q3: 3 * (11 + 1) / 4 = 36 / 4 = 9. Nilai Q3 adalah data ke-9, yaitu **35**.

- *Contoh Interpolasi:* Data: 5, 8, 10, 12, 15, 18, 20, 22 (n=8).
  - Posisi Q1: 1 * (8 + 1) / 4 = 9 / 4 = 2.25 (k=2, d=0.25).
  - Nilai Q1 = Data ke-2 + 0.25 * (Data ke-3 - Data ke-2)
  - Nilai Q1 = 8 + 0.25 * (10 - 8) = 8 + 0.25 * 2 = 8 + 0.5 = 8.5.
  - Posisi Q3: 3 * (8 + 1) / 4 = 27 / 4 = 6.75 (k=6, d=0.75).
  - Nilai Q3 = Data ke-6 + 0.75 * (Data ke-7 - Data ke-6)
  - Nilai Q3 = 18 + 0.75 * (20 - 18) = 18 + 0.75 * 2 = 18 + 1.5 = 19.5.

➢ Perhitungan Kuartil Data Berkelompok (*Grouped Data*)

Formula untuk mengestimasi kuartil data berkelompok mirip dengan median:
Qi = L_Qi + [ ( (i * n / 4) - F ) / f_Qi ] * c

(Levine et al., 2019)

Di mana:

- Qi = Kuartil ke-i (i=1, 2, atau 3)
- L_Qi = Tepi bawah kelas kuartil ke-i. Kelas Qi adalah kelas interval di mana posisi i*n/4 berada.
- n = Jumlah total frekuensi.
- F = Frekuensi kumulatif *sebelum* kelas kuartil ke-i.
- f_Qi = Frekuensi kelas kuartil ke-i.
- c = Panjang interval kelas.

**Contoh**

Menggunakan data Pendapatan Karyawan

| Pendapatan (Juta Rp) | Frekuensi (f) | Frekuensi Kumulatif (F Kum) |
|---|---|---|
| 5 - < 7 | 8 | 8 |
| 7 - < 9 | 12 | 20 |
| 9 - < 11 | 15 | 35 |
| 11 - < 13 | 10 | 45 |
| 13 - < 15 | 5 | 50 |
| **Total** | **n = 50** | |

- Hitung Q1 (i=1):
  - Posisi Q1 = 1 * 50 / 4 = 12.5. Kelas Q1 adalah 7 - < 9 (F Kum pertama > 12.5).
  - L_Q1=7, F=8, f_Q1=12, c=2.
  - Q1 = 7 + [ (12.5 - 8) / 12 ] * 2 = 7 + [ 4.5 / 12 ] * 2 = 7 + 0.375 * 2 = 7 + 0.75 = 7.75.
  - *Interpretasi Q1:* 25% karyawan berpendapatan kurang dari 7.75 Juta Rupiah.
- Hitung Q3 (i=3):

- Posisi Q3 = 3 * 50 / 4 = 37.5. Kelas Q3 adalah 11 - < 13 (F Kum pertama > 37.5).
- L_Q3=11, F=35, f_Q3=10, c=2.
- Q3 = 11 + [ (37.5 - 35) / 10 ] * 2 = 11 + [ 2.5 / 10 ] * 2 = 11 + 0.25 * 2 = 11 + 0.5 = 11.5.
- *Interpretasi Q3:* 75% karyawan berpendapatan kurang dari 11.5 Juta Rupiah (atau 25% berpendapatan lebih dari 11.5 Juta Rupiah).

➢ Jangkauan Antar Kuartil (*Interquartile Range* - IQR)

IQR adalah selisih antara kuartil ketiga dan kuartil pertama: IQR = Q3 - Q1.
IQR mengukur penyebaran (*spread*) dari 50% data yang berada di tengah distribusi. Keunggulan utama IQR adalah ketahanannya terhadap nilai ekstrem (*robust to outliers)*, menjadikannya ukuran penyebaran yang baik untuk data miring (Lind et al., 2021).

- *Contoh (Data Pendapatan):* IQR = 11.5 - 7.75 = **3.75** (Juta Rupiah). Ini berarti rentang pendapatan untuk 50% karyawan di bagian tengah adalah 3.75 Juta Rupiah.

➢ Aplikasi Kuartil

Kuartil dan IQR sering digunakan dalam pembuatan *Box Plot* (*Box-and-Whisker Plot*), sebuah visualisasi grafis yang efektif untuk menunjukkan pemusatan, penyebaran, dan kemiringan data, serta mendeteksi outlier (Keller, 2017). Mereka juga berguna untuk membandingkan distribusi antar kelompok dan memahami variasi dalam data.

### 3.6.2. Desil (*Deciles*)

- Definisi

  Desil adalah nilai-nilai yang membagi sekumpulan data yang telah diurutkan menjadi sepuluh bagian yang sama besar. Ada sembilan titik desil (D1, D2, ..., D9). Setiap bagian mewakili 10% dari data. Perhatikan bahwa D5 = Q2 = Median.

- **Perhitungan Desil**

  Metode perhitungannya analog dengan kuartil, hanya pembaginya yang berbeda.

  - Data Tunggal: Posisi(Di) = i * (n + 1) / 10 (i=1, ..., 9). Gunakan interpolasi jika perlu.
  - Data Berkelompok: Di = L_Di + [ ( (i * n / 10) - F ) / f_Di ] * c (Keller, 2017)

    (L_Di = Tepi bawah kelas desil ke-i, f_Di = frekuensi kelas desil ke-i).
  - *Contoh (Data Pendapatan, hitung D7):* n=50.
    - Posisi D7 = 7 * 50 / 10 = 35. Tepat jatuh di batas atas kelas 9-<11. Berarti kelas D7 adalah 9 - < 11. (Atau, jika F Kum = posisi, ambil batas atas kelas tsb. Jika > posisi, ambil kelas tsb).
    - L_D7=9, F=20, f_D7=15, c=2.
    - D7 = 9 + [ (35 - 20) / 15 ] * 2 = 9 + [ 15 / 15 ] * 2 = 9 + 1 * 2 = 11.
    - *Interpretasi D7:* 70% karyawan berpendapatan kurang dari 11 Juta Rupiah.

- Aplikasi Desil

  Desil berguna untuk mengelompokkan data ke dalam segmen 10%. Misalnya, dalam analisis pendapatan, D9 sering digunakan untuk

menentukan batas pendapatan bagi kelompok 10% teratas.

### 3.6.3 Persentil (*Percentiles*)

- Definisi
  Persentil adalah nilai-nilai yang membagi sekumpulan data yang telah diurutkan menjadi seratus bagian yang sama besar. Ada 99 titik persentil (P1, P2, ..., P99). Persentil ke-i (Pi) menunjukkan nilai di mana i% data berada di bawahnya. Hubungan penting: P25 = Q1, P50 = Median = Q2 = D5, P75 = Q3.
- Perhitungan Persentil
  Metode perhitungannya juga analog dengan kuartil dan desil.
  - Data Tunggal: Posisi(Pi) = i * (n + 1) / 100 (i=1, ..., 99). Gunakan interpolasi jika perlu.
  - Data Berkelompok: Pi = L_Pi + [ ( (i * n / 100) - F ) / f_Pi ] * c (Lind et al., 2021) (L_Pi = Tepi bawah kelas persentil ke-i, f_Pi = frekuensi kelas persentil ke-i).
  - *Contoh (Data Pendapatan, hitung P90):* n=50.
    - Posisi P90 = 90 * 50 / 100 = 45. Tepat jatuh di batas atas kelas 11-<13. Berarti kelas P90 adalah 11 - < 13.
    - L_P90=11, F=35, f_P90=10, c=2.
    - P90 = 11 + [ (45 - 35) / 10 ] * 2 = 11 + [ 10 / 10 ] * 2 = 11 + 1 * 2 = 13.
    - *Interpretasi P90:* 90% karyawan berpendapatan kurang dari 13 Juta Rupiah (atau 10% teratas berpendapatan 13 Juta Rupiah atau lebih).
- Aplikasi Persentil

Persentil sangat umum digunakan dalam berbagai bidang:

- Pendidikan: Skor tes standar sering dilaporkan dalam bentuk persentil (misal: "skor Anda berada di persentil ke-85", artinya skor Anda lebih tinggi dari 85% peserta tes lainnya).
- Kesehatan: Kurva pertumbuhan bayi dan anak menggunakan persentil untuk membandingkan berat dan tinggi badan dengan standar populasi.
- Keuangan: Ukuran risiko seperti Value at Risk (VaR) sering dinyatakan sebagai persentil kerugian (misal: VaR 1 hari 99% sebesar 1 Milyar Rupiah berarti ada 1% kemungkinan kerugian dalam sehari melebihi 1 Milyar Rupiah).
- Manajemen Kinerja: Menetapkan target atau benchmark berdasarkan persentil kinerja (misal: target penjualan adalah mencapai persentil ke-80 di antara semua cabang).

### 3.6.4 Rangkuman Ukuran Letak

Kuartil, desil, dan persentil (secara kolektif sering disebut *fractiles* atau *quantiles*) adalah ukuran statistik deskriptif yang menunjukkan posisi data relatif terhadap keseluruhan distribusi. Mereka melengkapi ukuran pemusatan data dengan memberikan wawasan tentang penyebaran dan peringkat data, yang sangat berguna dalam analisis ekonomi dan bisnis.

## DAFTAR PUSTAKA

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., Cochran, J. J., Fry, M. J., & Ohlmann, J. W. (2020). *Statistics for business and economics* .

Berenson, M. L., Levin, D. M., Szabat, K. A., & Stephan , D. F. (2020). *Basic Business Statistics: Concepts and Applications*. Pearson.com. https://www.pearson.com/en-us/subject-catalog/p/basic-business-statistics-concepts-and-applications/P200000006245/9780137400119?srsltid=AfmBOopwH-XQR6hRaPYBw6_LhyoI-OHS1ucGauqVcpTQmLUXs4rLZbqu

Berk, J. B., & Demarzo, P. M. (2020). *Corporate Finance* (5th ed.). Pearson.

BPS. (2021). *Survei Sosial Ekonomi Nasional (Susenas) September 2021 - Berita*. Bps.go.id; Badan Pusat Statistik Kabupaten Sumba Barat. https://sumbabaratkab.bps.go.id/id/news/2021/09/23/70/survei-sosial-ekonomi-nasional--susenas--september-2021.html

Brooks, C. (2019). *Introductory Econometrics for Finance* (4th ed.). Cambridge University Press.

Davenport, T. H., Harris, J. G., & Morison, R. (2017). *Analytics*

*at Work : Smarter decisions, Better Results*. Harvard Business Press.

Dessler, G. (2019). *Human resource management* (16th ed.). Harlow, England Pearson.

Galit Shmueli, Bruce, P. C., Gedeck, P., & Patel, N. R. (2019). *Data Mining for Business Analytics*. John Wiley & Sons.

Gravetter, F. J., & Wallnau, L. B. (2019). *Statistics for the behavioral sciences*. Cengage.

Grolemund, G., & Wickham, H. (2017). *R for Data Science*. Had.co.nz. https://r4ds.had.co.nz/

Hand, D. J. (2018). Statistical challenges of administrative and transaction data. *Journal of the Royal Statistical Society: Series a (Statistics in Society)*, *181*(3), 555–605. https://doi.org/10.1111/rssa.12315

Hartatik, Bernadetta Kwintiana, Titin Agustin Nengsih, Abdillah Baradja, & Terttiaavini. (2023, May 21). *DATA SCIENCE FOR BUSINESS (Pengantar dan Penerapan Berbagai Sektor)*. https://www.researchgate.net/publication/371732658_DATA_SCIENCE_FOR_BUSINESS_Pengantar_dan_Penerapan_Berbagai_Sektor

Healey, J. (2018). *A Tool for Social Research* . Amazon.com. https://www.amazon.com/Statistics-Social-Research-Analysis-MindTap/dp/0357371070

Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B., & Sethupathy, G. (2016, December 7). *The age of analytics: Competing in a data-driven world | McKinsey*. Www.mckinsey.com. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-age-of-analytics-competing-in-a-data-driven-world

Indonesia, B. (2020). *Laporan Kebijakan Moneter Triwulan I 2020*. Www.bi.go.id.

https://www.bi.go.id/id/publikasi/laporan/Pages/Laporan-Kebijakan-Moneter-Triwulan-I-2020.aspx

Keller, G. (2017). *Statistics for management and economics*. Cengage Learning.

Kotler, P., & Keller, K. L. (2018). *Marketing management* (15th ed.). Pearson.

Levine, D. M., Stephan, D. F., & Szabat , K. A. (2019). *Business Statistics: A First Course*. Pearson.com. https://www.pearson.com/en-us/subject-catalog/p/business-statistics-a-first-course/P200000006243/9780136880974?srsltid=AfmBOor_kxvzpmQLWjcxTROPHTFYk0-AwGMFEBLvrbxNakPspactLNvD

Levine, D., & Szabat, K. (2017). *Statistics for Managers Using Microsoft Excel: Levine, David, Stephan, David, Szabat, Kathryn: 9780134173054: Amazon.com: Books*. Amazon.com. https://www.amazon.com/Statistics-Managers-Using-Microsoft-Excel/dp/0134173058

Lind, D. A., Marchal, W. G., & Samuel Adam Wathen. (2021). *Statistical techniques in business & economics*. Mcgraw-Hill/Irwin.

Mayer-Schönberger, V., & Cukier, K. (2018). *Big Data: A Revolution That Will Transform How We Live, Work and Think*. John Murray.

McAfee, A., & Brynjolfsson, E. (2020). *Amazon.com: Machine, Platform, Crowd: Harnessing Our Digital Future: 9780393254297: McAfee, Andrew, Brynjolfsson, Erik: Books*. Amazon.com. https://www.amazon.com/Machine-Platform-Crowd-Harnessing-Digital/dp/0393254291

Montgomery, D. C., & Runger, G. C. (2018). *Applied statistics and probability for engineers*. Wiley.

N Gregory Mankiw. (2020). *Principles of economics*. Cengage

Learning Asia Pte Ltd.

Nate Silver. (2016). *The Signal and the Noise : Why So Many Predictions Fail-- But Some Don't*. Penguin Books.

OJK. (2020). *Statistik Pasar Modal Indonesia*. Ojk.go.id. https://ojk.go.id/id/kanal/pasar-modal/data-dan-statistik/statistik-pasar-modal/default.aspx

Porter, T. M. (2018). *The Rise of Statistical Thinking, 1820–1900*. Princeton.edu; Princeton University Press. https://press.princeton.edu/books/paperback/9780691208428/the-rise-of-statistical-thinking-1820-1900?srsltid=AfmBOopb4zTB1j5fAftExwPwMhFVnb1-qysq2imbhbhsOysBVRvqwDwS

Stevenson, W. J. (2021). *Operations Management* (14th ed.). Mcgraw-Hill Education.

Stigler, S. M. (2016). *The Seven Pillars of Statistical Wisdom — Harvard University Press*. Harvard University Press. https://www.hup.harvard.edu/books/9780674088917

Todaro, M. P., & Smith, S. C. (2020). *Economic Development*. Addison-Wesley Longman.

Weiss, N. (2019). *Introductory Statistics*. Pearson.

Wooldridge, J. M. (2020). *Introductory Econometrics: a modern approach.* (7th ed.). Cengage Learning.

# BAB 4
# UKURAN PENYEBARAN DATA

**Oleh Toto Hermawan**

## 4.1 Pendahuluan

Dalam upaya membandingkan beberapa kumpulan data, penggunaan ukuran pemusatan saja tidak cukup untuk memberikan representasi yang komprehensif terhadap karakteristik data. Bahkan, dalam beberapa kasus, hal ini dapat menghasilkan interpretasi yang keliru (Kustituanto & Badrudin, 1994:94). Oleh karena itu, ukuran penyebaran berperan sebagai pelengkap ukuran pemusatan dalam mendeskripsikan distribusi data secara lebih akurat. Ukuran penyebaran menggambarkan sejauh mana data menyimpang dari nilai rata-ratanya, di mana semakin kecil nilai penyebaran, semakin terpusat data tersebut, dan sebaliknya.

## 4.2 Ukuran Penyebaran data

Ukuran penyebaran data merupakan aspek fundamental dalam statistika yang menunjukkan tingkat variasi atau penyebaran suatu kumpulan data dari nilai rata-ratanya. Konsep ini berperan penting dalam memahami konsistensi serta variabilitas data, yang esensial untuk mengidentifikasi karakteristik utama dari data tersebut. Dengan mengetahui ukuran penyebaran, peneliti dan analis data dapat menarik kesimpulan yang lebih akurat,

mendeteksi keberadaan nilai pencilan (*outlier*), serta menilai keandalan hasil analisis. Dalam kajian mengenai ukuran penyebaran, beberapa materi utama yang dibahas meliputi jangkauan (*range*), Simpangan Rata-rata, varians, dan Standar Deviasi.

### 4.2.1 Rentang (*Range*)

Rentang (*range*) merupakan metode paling sederhana dalam mengukur tingkat penyebaran suatu data. Rentang didefinisikan sebagai selisih antara nilai tertinggi dan nilai terendah dalam suatu kumpulan data. Secara matematis, jangkauan dirumuskan sebagai:

$$R = X_{maks} - X_{\min}$$

Keterangan :

R = range

$X_{maks}$ = Skor terbesar

$X_{\min}$ = Skor terkecil

Makna dari nilai rentang adalah Semakin kecil nilai Rentangnya maka penyebaran menunjukkan karakteristik data yang lebih baik, karena hal ini mengindikasikan bahwa data lebih terpusat di sekitar nilai tengah dan memiliki tingkat konsistensi yang lebih tinggi.

**Contoh :**

1. Tentukan range/ jangkuan dari data : 10,6,8,2,4

Jawab : R = Xmaks – Xmin = 10 – 2 = 8

2. Diberikan data distribusi berkelompok sebagai berikut :

| nilai | Frekuensi |
|---|---|
| 3 - 5 | 3 |
| 6 - 8 | 6 |
| 9 - 11 | 16 |
| 12 - 14 | 8 |
| 15 - 17 | 7 |
| 18 - 20 | 10 |

Jawab :

*Catatan : Rentang pada data berkelompok adalah selisih antara batas atas dari kelas tertinggi dengan batas bawah dari kelas terendah*

Maka

Tepi bawah kelas terendah = 2,5

Tepi atas kelas tertinggi = 20,5

*Rentang* 20,50 – 2,50 = 18

**Kelemahan**

Rentang (*range*) mengukur tingkat penyebaran keseluruhan dalam suatu kumpulan data. Meskipun perhitungannya sederhana, Rentang memiliki keterbatasan dalam merepresentasikan distribusi data secara menyeluruh. Ukuran ini tidak memberikan informasi mengenai bagaimana data tersebar antara nilai minimum dan maksimum, sehingga kurang efektif dalam menggambarkan karakteristik data secara komprehensif jika digunakan sebagai satu-satunya ukuran penyebaran.

### 4.2.2 Simpangan Rata-rata ( *Mean Deviation*)

Simpangan rata-rata merupakan salah satu ukuran penyebaran data, sebagaimana halnya Rentang. Ukuran ini digunakan untuk mengukur sejauh mana setiap nilai dalam suatu kumpulan data menyimpang dari rata-rata sebenarnya, sehingga memberikan gambaran mengenai tingkat variabilitas data.

Simpangan rata-rata merupakan alat yang sangat bermanfaat dalam analisis statistik, terutama dalam menilai kestabilan suatu data. Ukuran ini memberikan gambaran tentang tingkat konsistensi data dalam suatu distribusi. Jika simpangan rata-rata bernilai kecil, maka data dapat dikatakan stabil dan memiliki variasi yang rendah. Sebaliknya, jika simpangan rata-rata besar, data menunjukkan tingkat ketidakstabilan atau variabilitas yang tinggi. Informasi ini memiliki berbagai aplikasi, seperti dalam pengendalian kualitas produk maupun penelitian ilmiah, di mana konsistensi data sangat penting untuk validitas hasil analisis.

1. Data Tunggal

   Secara matematis, Simpangan Rata-rata data tunggal dirumuskan sebagai:

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

   Keterangan :

   SR = Simpangan Rata-rata

   $\bar{x}$ = Rata rata

   $x_i$ = nilai data ke-i

   n = banyak data

**Contoh :**
Nilai ulangan matamatika dari 6 siswa adalah :7,5,6,3,8,7.
Tentukan simpangan rata-ratanya!

Jawab :
Langkah 1 cari nilai rata rata data diatas

$$\bar{x} = \frac{7+5+6+3+8+7}{6}$$

Langkah 2 mencari SR

$$SR = \frac{|7-6|+|5-6|+|6-6|+|3-6|+|8-6|+|7-6|}{6}$$

$$SR = \frac{1+1+0+3+2+1}{6}$$

$$SR = \frac{8}{6}$$

2. Data Berkelompok
Secara matematis, Simpangan Rata-rata data berkelompok dirumuskan sebagai:

$$SR = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i}$$

Keterangan :

SR = Simpangan Rata-rata

$\bar{x}$ = Rata rata

$x_i$ = nilai tengah kelas ke-i

$f_i$ = frekuensi kelas ke-i

**Contoh :**

Tentukan nilai simpangan rata-rata dari table berikut!

| nilai | Frekuensi |
|---|---|
| 31 - 35 | 1 |
| 36 - 40 | 2 |
| 41 – 45 | 3 |
| 46 – 50 | 4 |
| Jumlah | 10 |

Jawab :

Langkah 1 cari nilai rata rata data diatas

| nilai | f | $x_i$ | $f.x_i$ |
|---|---|---|---|
| 31 - 35 | 1 | 33 | 33 |
| 36 - 40 | 2 | 38 | 76 |
| 41 – 45 | 3 | 43 | 129 |
| 46 – 50 | 4 | 48 | 192 |
| Jumlah | 10 | | 430 |

Maka

$x_1$ = nilai tengah kelas ke-1= $\frac{35+31}{2} = \frac{66}{2} = 33$

$x_2$ = nilai tengah kelas ke-2= $\frac{36+40}{2} = \frac{76}{2} = 38$

$x_3$ = nilai tengah kelas ke-3= $\frac{41+45}{2} = \frac{86}{2} = 43$

$x_4$ = nilai tengah kelas ke-4= $\frac{46+50}{2} = \frac{96}{2} = 48$

Akibatnya :

$f_1.x_1 = 1x\ 33 = 33$

$f_2.x_2 = 2x\ 38 = 76$

$f_3.x_3 = 3x\ 43 = 129$

$f_4.x_4 = 4x\ 48 = 192$

Sehingga

$$\bar{x} = \frac{430}{10} = 43$$

Langkah 2 mencari SR dengan mengunakan tabel bantu

| nilai | f | $x_i$ | $\lvert x_i - \bar{x}\rvert$ | $f_i\lvert x_i - \bar{x}\rvert$ |
|---|---|---|---|---|
| 31 - 35 | 1 | 33 | \| 33 – 43 \| =10 | 10 |
| 36 - 40 | 2 | 38 | \| 38 – 43 \| = 5 | 10 |
| 41 – 45 | 3 | 43 | \| 43 – 43 \| = 0 | 0 |
| 46 – 50 | 4 | 48 | \| 48 – 43 \| =5 | 20 |
| Jumlah | 10 | | | 40 |

Sehingga dengan mudah kita dapatkan SR sebagai berikut :

$$SR = \frac{40}{10} = 4$$

**Kelemahan**

Simpangan rata-rata mempertimbangkan seluruh nilai dalam data untuk mengukur tingkat penyebaran. Namun, karena perhitungannya menggunakan nilai absolut, metode ini tidak memberikan informasi mengenai arah penyebaran data, apakah lebih condong ke atas atau ke bawah dari rata-rata.

### 4.2.3 Varians

Variansi adalah ukuran yang menunjukkan seberapa jauh data dalam suatu himpunan tersebar dari nilai rata-ratanya. Semakin besar variansi, semakin besar variasi atau perbedaan antar nilai dalam data. Sebaliknya, variansi yang kecil menunjukkan bahwa data cenderung berkelompok di sekitar rata-rata, mencerminkan tingkat konsistensi yang lebih tinggi

dalam distribusi data. Varians untuk data populasi disimbolkan $\sigma^2$ dan untuk data sampel disimbolkan dengan $s^2$

1. Data Tunggal

   Secara matematis,Varians data tunggal dirumuskan sebagai:

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

   Keterangan :

   $s^2$ = Varians

   $\bar{x}$ = Rata rata

   n = banyak data

   **Contoh :**

   Diketahui Nilai ulangan matamatika dari 6 siswa adalah : 7 , 5 , 6, 3, 8, 7. Tentukan varians data diatas !

   Jawab :

   Langkah 1 cari nilai rata rata data diatas

$$\bar{x} = \frac{7+5+6+3+8+7}{6}$$

   Langkah 2 mencari varians

$$s^2 = \frac{(7-6)^2 + (5-6)^2 + (6-6)^2 + (3-6)^2 + (8-6)^2 + (7-6)^2}{6}$$

$$s^2 = \frac{1^2 + (-1)^2 + 0 + (-3)^2 + 2^2 + 1^2}{6}$$

$$s^2 = \frac{1+1+0+9+4+1}{6}$$

$$s^2 = \frac{16}{6}$$

2. Data Kelompok

Secara matematis,Varians data kelompok dirumuskan sebagai:

$$s^2 = \frac{\sum f_i \left(x_i - \bar{x}\right)^2}{\sum f_i}$$

Keterangan :

$s^2$ = varians

$\bar{x}$ = Rata rata

$x_i$ = nilai tengah kelas ke-i

$f_i$ = frekuensi kelas ke-i

**Contoh :**

Tentukan nilai varians dari table berikut!

| nilai | Frekuensi |
|---|---|
| 31 - 35 | 1 |
| 36 - 40 | 2 |
| 41 – 45 | 3 |
| 46 – 50 | 4 |
| Jumlah | 10 |

Jawab :

Langkah 1 cari nilai rata rata data diatas

| nilai | f | $x_i$ | $f.x_i$ |
|---|---|---|---|
| 31 - 35 | 1 | 33 | 33 |
| 36 - 40 | 2 | 38 | 76 |
| 41 – 45 | 3 | 43 | 129 |
| 46 – 50 | 4 | 48 | 192 |
| Jumlah | 10 | | 430 |

Sehingga

$\bar{x} = \frac{430}{10} = 43$ *(cara sama seperti contoh diatas)*

Langkah 2 cari varians mengunakan tabel bantu

| nilai | f | $x_i$ | $(x_i - \bar{x})^2$ | $f_i(x_i - \bar{x})^2$ |
|---|---|---|---|---|
| 31 - 35 | 1 | 33 | 100 | 100 |
| 36 - 40 | 2 | 38 | 25 | 50 |
| 41 – 45 | 3 | 43 | 0 | 0 |
| 46 – 50 | 4 | 48 | 25 | 100 |
| Jumlah | 10 | | | 250 |

$$s^2 = \frac{250}{10} = 25$$

**Kelemahan**

Perlu diperhatikan bahwa penggunaan variansi cenderung menghasilkan nilai dispersi yang lebih besar dibandingkan dengan simpangan rata-rata. Hal ini disebabkan oleh perhitungan variansi yang menggunakan kuadrat selisih setiap nilai data terhadap rata-rata, sehingga meningkatkan nilai penyebaran secara signifikan. Akibatnya, variansi kurang efektif sebagai ukuran dispersi yang langsung menggambarkan sebaran data. Kelemahan utama variansi terletak pada sifat perhitungannya yang berbasis kuadrat, sedangkan penyebaran data secara alami lebih bersifat linier. Oleh karena itu, dalam analisis data, variansi jarang digunakan secara langsung sebagai indikator utama penyebaran. Meskipun demikian, variansi tetap memiliki keunggulan karena mempertimbangkan seluruh selisih nilai data terhadap rata-rata, sehingga tetap relevan dalam berbagai analisis statistik.

### 4.2.4 Standar Deviasi

Standar Deviasi (Simpangan baku) merupakan ukuran statistik yang menggambarkan tingkat variasi atau penyebaran suatu kumpulan data terhadap nilai rata-ratanya. Nilai ini diperoleh dengan menghitung akar kuadrat dari variansi, sehingga memungkinkan interpretasi dalam satuan yang sama dengan data aslinya. Semakin kecil simpangan baku, semakin dekat data dengan nilai rata-rata, menunjukkan distribusi yang lebih terpusat. Sebaliknya, simpangan baku yang lebih besar menandakan bahwa data tersebar lebih luas dari nilai rata-rata.

1. Data Tunggal

   Secara matematis, Standar Deviasi data tunggal dirumuskan sebagai:

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

   Keterangan :

   S = Standar Deviasi

   **Contoh :**

   Diketahui Nilai ulangan matamatika dari 6 siswa adalah : 7 , 5 , 6, 3, 8, 7. Tentukan Standar Deviasi data diatas !

   **Jawab :**

   Dari contoh diatas sudah dibahas cara mencari $s^2 = \frac{16}{6}$

   Sehingga dengan mudah kita bisa mendapatkan Standar Deviasi dengan cara $s = \sqrt{s^2}$

Akibatnya $s = \sqrt{16/6}$

Jadi Standar Deviasi(Simpangan baku) adalah $s = \sqrt{16/6}$

2. Data Kelompok

Secara matematis, Standar Deviasi data Kelompok dirumuskan sebagai:

$$s = \sqrt{\frac{\sum f_i \left(x_i - \bar{x}\right)^2}{\sum f_i}}$$

Keterangan :

S = Standar Deviasi

**Contoh :**

Tentukan nilai varians dari *table* berikut!

| nilai | Frekuensi |
|---|---|
| 31 - 35 | 1 |
| 36 - 40 | 2 |
| 41 – 45 | 3 |
| 46 – 50 | 4 |
| Jumlah | 10 |

**Jawab :**

Dari contoh diatas sudah dibahas cara mencari $s^2 = 25$

Sehingga dengan mudah kita bisa mendapatkan Standar Deviasi dengan cara $s = \sqrt{s^2}$

Akibatnya $s = \sqrt{25}$

Jadi Standar Deviasi(Simpangan baku) adalah $s = 5$

### 4.2.5 Ukuran Penyebaran Relatif

1. Koefisien Jangkauan

   Koefisien Jangkauan adalah ukuran relatif dari penyebaran data yang dihitung dengan membandingkan selisih antara nilai maksimum dan minimum terhadap jumlahnya. Koefisien ini digunakan untuk menilai tingkat variasi dalam suatu kumpulan data secara relatif. Secara matematis, Koefisien Jangkauan dirumuskan sebagai:

$$KJ = \frac{X_{maks} - X_{\min}}{X_{maks} + X_{\min}} x100\%$$

   Keterangan :

   KJ = range

   $X_{maks}$ = Skor terbesar

   $X_{\min}$ = Skor terkecil

   Interpretasi dari KJ adalah Nilai yang lebih tinggi menunjukkan variasi data yang lebih besar, sedangkan nilai yang lebih kecil menunjukkan data yang lebih homogen atau terpusat.

   **Contoh :**
   Tentukan KJ dari data : 10,6,8,2,4
   **Jawab :**

   $X_{maks}$ = 10

   $X_{\min}$ = 2

   Maka

$$KJ = \frac{10-2}{10+2} x100\%$$

$$KJ = \frac{8}{12} x100\%$$
$$KJ = 66,67\%$$

2. Koefisien Deviasi Rata-rata

Koefisien Deviasi Rata-rata adalah ukuran relatif dari penyebaran data yang dinormalisasi terhadap rata-rata hitung (mean). Koefisien ini digunakan untuk membandingkan tingkat variasi antara kumpulan data yang memiliki skala berbeda. Secara matematis, Koefisien Deviasi Rata-rata dirumuskan sebagai:

$$KDR = \frac{SR}{\bar{X}} x100\%$$

Keterangan :

KDR = Koefisien Deviasi Rata-rata

SR = Deviasi Rata-rata

$\bar{X}$ = rata rata hitung

**Contoh :**

Diketahui rata rata hitung suatu data adalah 10. Selain itu simpangan rata rata data tersebut adalah 4. Hitunglah Koefisien Deviasi Rata-rata data tersebut.

**Jawab :**

$$KDR = \frac{4}{10} x100\%$$
$$KDR = 40\%$$

Koefisien Deviasi Rata-rata adalah 40 %

3. Koefisien Standar Deviasi

Koefisien Standar Deviasi, atau dikenal sebagai Koefisien Variasi (*Coefficient of Variation,* CV), adalah ukuran relatif dari penyebaran data yang dinyatakan dalam bentuk persentase terhadap rata-rata hitung. Koefisien ini digunakan untuk membandingkan tingkat variasi antara kumpulan data yang memiliki satuan atau skala yang berbeda. Secara matematis, Koefisien Standar Deviasi dirumuskan sebagai:

$$CV = \frac{s}{\bar{X}} x100\%$$

Keterangan :

CV = Koefisien Standar Deviasi

s = standar deviasi

$\bar{X}$ = rata rata hitung

**Interpretasi:**

1) CV rendah menunjukkan bahwa variasi data relatif kecil dibandingkan dengan rata-rata, sehingga data lebih homogen.
2) CV ting**gi** menunjukkan bahwa data lebih tersebar, sehingga variasinya lebih besar dibandingkan dengan rata-rata.

Koefisien standar deviasi sering digunakan dalam analisis ekonomi, keuangan, dan penelitian ilmiah untuk membandingkan variabilitas antar kelompok data dengan skala yang berbeda.

**Contoh :**

Diketahui rata rata hitung suatu data adalah 10. Selain itu standar deviasi data tersebut adalah 2,5 . Hitunglah Koefisien standar deviasi data tersebut.

Jawab:

$$CV = \frac{2,5}{10} x100\%$$

$$CV = 25\%$$

Koefisien Koefisien standar deviasi adalah 25 %

## 4.3 Kesimpulan

Ukuran penyebaran menggambarkan sejauh mana data menyimpang dari nilai rata-ratanya, di mana semakin kecil nilai penyebaran, semakin terpusat data tersebut, dan sebaliknya.

a. jangkauan (*range*),
b. Simpangan Rata-rata,
c. varians , dan
d. Standar Deviasi.

## 4.4 Latihan

| nilai | Frekuensi |
|---|---|
| 31 - 35 | 1 |
| 36 - 40 | 5 |
| 41 – 45 | 4 |
| 46 – 50 | 6 |
| 51 – 55 | 4 |
| 56 – 60 | 5 |
| Jumlah | 25 |

Carilah nilai dari

a. jangkauan (*range*),
b. Simpangan Rata-rata,
c. varians , dan
d. Standar Deviasi.

## DAFTAR PUSTAKA

Kustituanto, B., & Badrudin, R. (1994). Statistika 1 (Deskriptif). Penerbit Gunadarma

Lee William, Carlson; Paul Newbold and ThorneBetty, 2013. Statistics for Business and Economics. Pearson Education.

Sudjana.2002.Metoda Statistika.Bandung : Tarsito

Sugiyono. 2016. Statistika Untuk Penelitian. Bandung: Alfabeta

# BAB 5
# UKURAN LETAK DATA

**Oleh Muhammad Zulkarnain**

## 5.1 Pengertian Ukuran Letak data

Ukuran letak data adalah suatu ukuran statistik yang digunakan untuk menentukan posisi suatu nilai dalam distribusi sekumpulan data. Ukuran ini berguna dalam analisis data untuk memahami persebaran dan perbandingan nilai dalam suatu kumpulan data (Sudjana, 1996). Ukuran letak data sangat berguna dalam berbagai analisis statistik, seperti menentukan distribusi nilai ujian, membandingkan kinerja individu dalam suatu kelompok, serta memahami variabilitas data dalam suatu populasi (Walpole, 1992).

## 5.2 Jenis-Jenis Ukuran Letak Data

Secara umum, ukuran letak data terdiri dari beberapa jenis, yaitu kuartil, desil, dan persentil (Walpole, 1992; Sudjana, 1996). Masing-masing ukuran ini membagi data menjadi bagian-bagian yang sama besar, yaitu :

### 5.2.1 Kuartil (Q)

Kuartil adalah ukuran letak dalam statistik yang membagi data yang telah diurutkan menjadi empat bagian yang sama besar. Setiap bagian mewakili 25%

dari keseluruhan data, sehingga kuartil digunakan untuk memahami distribusi data dalam suatu kelompok (Sudjana, 1996). Kuartil sering digunakan dalam berbagai bidang seperti ekonomi, bisnis, dan pendidikan untuk menganalisis persebaran data dan membandingkan nilai dalam suatu populasi.

Kuartil terdiri dari tiga nilai penting yang membagi data menjadi empat bagian, yaitu (Walpole, 1992):

a. Kuartil pertama (Q1) atau Kuartil Atas

   Kuartil pertama (Q1) adalah nilai yang membagi 25% data terbawah dari keseluruhan data, artinya, 25% data berada di bawah Q1, dan 75% data berada di atasnya.

b. Kuartil kedua (Q2) atau Kuartil Tengah (Median)

   Kuartil kedua (Q2) adalah nilai tengah dari seluruh data yang telah diurutkan, kuartil kedua ini juga dikenal sebagai median, karena membagi data menjadi dua bagian yang sama besar (50% di bawah dan 50% di atas).

c. Kuartil ketiga (Q3) atau Kuartil Bawah

   Kuartil ketiga (Q3) adalah nilai yang membagi 75% data terbawah dari keseluruhan data, artinya, 75% data berada di bawah Q3, dan 25% sisanya berada di atasnya.

Dalam statistik, Kuartil sering digunakan dalam berbagai bidang seperti ekonomi, bisnis, dan pendidikan untuk menganalisis persebaran data dan membandingkan nilai dalam suatu populasi.

### 5.2.2 Desil (D)

Desil adalah salah satu ukuran pemusatan data yang membagi suatu kumpulan data yang telah diurutkan menjadi sepuluh bagian yang sama besar. Setiap bagian disebut sebagai desil ke-1 hingga desil ke-9, yang berfungsi untuk menggambarkan distribusi data dalam kelompok tertentu (Trihendradi, 2019).

Dalam statistik, desil sering digunakan untuk menganalisis distribusi data, khususnya dalam bidang ekonomi, pendidikan, dan penelitian sosial. Misalnya, dalam analisis pendapatan, desil digunakan untuk mengelompokkan pendapatan penduduk ke dalam sepuluh bagian, sehingga dapat diketahui seberapa besar ketimpangan ekonomi dalam suatu populasi (Sudjana, 1996).

### 5.2.3 Persentil (P)

Persentil adalah ukuran statistik yang membagi data yang telah diurutkan menjadi 100 bagian yang sama besar. Persentil sering digunakan dalam analisis ekonomi dan bisnis untuk memahami distribusi data, seperti pendapatan, harga saham, atau kinerja perusahaan. Setiap bagian disebut sebagai persentil pertama hingga persentil ke-99, yang menunjukkan posisi relatif suatu nilai dalam distribusi data (Walpole, 1992).

Dalam statistik, Persentil sering digunakan untuk menilai posisi individu dalam suatu kelompok (misalnya, skor ujian dalam pendidikan), menganalisis distribusi pendapatan dalam ekonomi, menentukan pertumbuhan bayi dalam bidang Kesehatan (Walpole, 1992).

Jika seseorang berada pada persentil ke-75 dalam skor ujian, itu berarti 75% peserta ujian memiliki skor lebih rendah, dan 25% memiliki skor lebih tinggi.

## 5.3 Rumus dan Contoh Perhitungan Ukuran Letak Data

Rumus dan contoh perhitungan ukuran letak data terbagi menjadi rumus data tunggal dan rumus data berkelompok. Data tunggal adalah data yang disusun secara tunggal, tidak dalam bentuk interval. Sedangkan data berkelompok, adalah kumpulan data yang ditulis dalam bentuk interval.

### 5.3.1 Rumus Kuartil (Q)

**a. Rumus Kuartil untuk Data Tunggal**

Jika data berjumlah **n**, maka kuartil dihitung dengan rumus (Triola, 2014):

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

Di mana:

$Q_k$ = Kuartil ke-k (Q1, Q2, atau Q3),
$k$ = Nomor kuartil (1, 2, atau 3),
$n$ = Jumlah data.

**Contoh Perhitungan** :

Seorang analis pasar ingin mengetahui Q2 (Median) dan Q3 (Kuartil ke-3) dari harga saham suatu perusahaan dalam 10 hari terakhir (dalam ribuan rupiah) dengan data harga saham setelah diurutkan sebagai berikut :

50, 52, 55, 58, 60, 63, 65, 70, 75, 80

Langkah Perhitungan :
Dari data tersebut, diketahui bahwa $n = 10$, maka :

**Menghitung Q2 (Median)**

Untuk mengetahui letak posisi median (Q2), maka dilakukan perhitungan sebagai berikut :

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

$$Q2 = \frac{2(10+1)}{4} = \frac{2(11)}{4} = \frac{22}{4} = 5{,}5$$

Hasil 5,5 menunjukkan bahwa posisi median (Q2) terletak di antara data ke-5 dan ke-6, yaitu 60 (X5) dan 63 (X6).

50, 52, 55, 58, 60, 63, 65, 70, 75, 80

Hasil perhitungan ini menunjukkan posisi dalam bentuk pecahan, maka dilakukan interpolasi atau rata-rata antara dua nilai terdekat yang digunakan.

Cara Interpolasi, sebagai berikut :

$$Q2 = X5 + 0{,}5(X6 - X5)$$

$(Catatan : 0{,}5\ berasal\ dari\ pemisahan\ 5{,}5\ (5 + 0{,}5)$

$$Q2 = 60 + 0{,}5(63 - 60) = 60 + 0{,}5(3)$$

$$Q2 = 60 + 1{,}5 = 61{,}5$$

Maka Q2 (Median) terletak di posisi 61,5.
Jadi Kuartil ke-2 (median) dari harga saham perusahaan adalah Rp 61.500,-. Artinya, 50 % dari

harga saham perusahaan tersebut berada di bawah Rp 61.500,- dan sisanya di atas angka tersebut.

**Menghitung Q3 (Kuartil 3)**

Untuk mengetahui letak posisi Kuartil 3 (Q3), maka dilakukan perhitungan sebagai berikut :

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

$$Q3 = \frac{3(10+1)}{4} = \frac{3(11)}{4} = \frac{33}{4} = 8{,}25$$

Hasil 8.25 berarti nilai kuartil terletak di antara data ke-8 dan ke-9, yaitu 70 (X8) dan 75 (X9).

50, 52, 55, 58, 60, 63, 65, 70, 75, 80

Karena hasil Q3 dalam bentuk pecahan, maka dilakukan Interpolasi, sebagai berikut :

$$Q3 = X8 + 0{,}25(X9 - X8)$$

($Catatan$ : 0,25 *berasal dari pemisahan* 8,25 (8 + 0,25)

$$Q3 = 70 + 0{,}25(75 - 70) = 60 + 0{,}25(5)$$

$$Q2 = 70 + 1{,}25 = 71{,}25$$

Maka Kuartil ke-3 (Q3) terletak di posisi 71,25.

Jadi Q3 dari harga saham perusahaan adalah Rp 71.250,-. Artinya, 75 % dari harga saham Perusahaan tersebut berada di bawah Rp 71.250,- dan sisanya di atas angka tersebut.

**b. Rumus Kuartil untuk Data Kelompok**

Jika data berbentuk distribusi frekuensi, kuartil dihitung menggunakan rumus (Triola, 2014):

$$Qk = L + \left( \frac{\frac{kN}{4} - F}{f} \right) c$$

dengan:

$L$ = Batas bawah kelas kuartil.
$N$ = Jumlah total frekuensi.
$F$ = Jumlah frekuensi sebelum kelas kuartil.
$f$ = Frekuensi kelas kuartil.
$c$ = Panjang interval kelas.

**Contoh Perhitungan** :

Seorang analis ekonomi sedang meneliti pendapatan bulanan karyawan di sebuah perusahaan. Data pendapatan bulanan (dalam juta rupiah) dikelompokkan dalam tabel distribusi frekuensi sebagai berikut :

| Kelas Interval | Frekuensi (f) |
|---|---|
| 2-4 | 5 |
| 5-7 | 8 |
| 8-10 | 12 |
| 11-13 | 10 |
| 14-16 | 5 |

Tentukanlah nilai dari Q1, Q2, dan Q3 dari pendapatan bulanan karyawan tersebut.

**Langkah Perhitungan** :
Dari tabel tersebut, diketahui total frekuensi (N) adalah :

$$N=5+8+12+10+5=40$$

**Menghitung Kuartil 1 (Q1)**

Langkah-Langkah :

1. Tentukan posisi Q1
   Untuk mengetahui letak posisi Kuartil 1 (Q1), maka dilakukan perhitungan sebagai berikut :

$$Q1 = \frac{1}{4}N = \frac{1}{4}.40 = 10$$

   Jadi, kuartil pertama (Q1) berada pada data ke-10.

2. Tentukan kelas Q1
   Untuk mengetahui kelas Q1, maka dilakukan penjumlahan frekuensi sebelum kelas kuartil , sebagai berikut :

| Kelas Interval | Frekuensi (f) | Frekuensi Kumulatif (F) |
|---|---|---|
| 2-4 | 5 | 5 |
| 5-7 | 8 | 13 |
| 8-10 | 12 | 25 |
| 11-13 | 10 | 35 |
| 14-16 | 5 | 40 |

Data ke-10 (Q1) berada di kelas 5-7, karena frekuensi kumulatif sebelumnya $(5) < 10 \leq F\ saat\ ini\ (13)$.

3. Gunakan rumus kuartil data kelompok

$$Qk = L + \left( \frac{\frac{kN}{4} - F}{f} \right) c$$

Untuk Q1 ($k = 1$) :

- L = 4,5 (batas bawah kelas Q1 adalah 5-0,5)
- N = 40
- F = 5 (frekuensi kumulatif sebelum kelas Q1)
- f = 8 (frekuensi kelas Q1)
- c = 2 (panjang kelas)

maka :

$$Q1 = 4{,}5 + \left( \frac{\frac{1.40}{4} - 5}{8} \right) . 2$$

$$Q1 = 4{,}5 + \left( \frac{\frac{40}{4} - 5}{8} \right) . 2$$

$$Q1 = 4{,}5 + \left( \frac{10 - 5}{8} \right) . 2$$

$$Q1 = 4{,}5 + \left( \frac{10}{8} \right) = 4{,}5 - 1{,}25 = 5{,}75$$

Dari perhitungan di atas diperoleh, diperoleh kuartil pertama (Q1) dari pendapatan bulanan karyawan tersebut adalah Rp 5.750.000,-

**Menghitung Kuartil 2 atau Median (Q2)**

Langkah-Langkah :

1. Tentukan posisi Q2
   Untuk mengetahui letak posisi Kuartil 2 (Q2), maka dilakukan perhitungan sebagai berikut :

$$Q2 = \frac{2}{4}N = \frac{2}{4}.40 = 20$$

Jadi, kuartil kedua (Q2) berada pada data ke-20.

2. Tentukan kelas Q2

   Untuk mengetahui kelas Q2, maka dilakukan penjumlahan frekuensi sebelum kelas kuartil , sebagai berikut :

| Kelas Interval | Frekuensi (f) | Frekuensi Kumulatif (F) |
|---|---|---|
| 2-4 | 5 | 5 |
| 5-7 | 8 | 13 |
| 8-10 | 12 | 25 |
| 11-13 | 10 | 35 |
| 14-16 | 5 | 40 |

   Data ke-20 (Q2) berada di kelas 8-10, karena frekuensi kumulatif sebelumnya $(13) < 20 \leq F\ saat\ ini\ (25)$.

3. Gunakan rumus kuartil data kelompok

$$Qk = L + \left( \frac{\frac{kN}{4} - F}{f} \right) c$$

   Untuk Q2 $(k = 2)$ :

   - L = 7,5 (batas bawah kelas Q2 adalah 8-0,5)
   - N = 40
   - F = 13 (frekuensi kumulatif sebelum kelas Q2)

- f = 12 (frekuensi kelas Q2)
- c = 2 (panjang kelas)

maka :

$$Q2 = 7{,}5 + \left(\frac{\frac{2.40}{4} - 13}{12}\right).2$$

$$Q2 = 7{,}5 + \left(\frac{\frac{80}{4} - 13}{12}\right).2$$

$$Q2 = 7{,}5 + \left(\frac{20 - 13}{12}\right).2$$

$$Q2 = 7{,}5 + \left(\frac{7}{12}\right).2$$

$$Q2 = 7{,}5 + \left(\frac{14}{12}\right) = 7{,}5 + 1{,}17 = 8{,}67$$

Dari perhitungan di atas diperoleh, diperoleh kuartil kedua atau median (Q2) dari pendapatan bulanan karyawan tersebut adalah Rp 8.670.000,-

**Menghitung Kuartil 3 (Q3)**

Langkah-Langkah :

1. Tentukan posisi Q3
   Untuk mengetahui letak posisi Kuartil 3 (Q3), maka dilakukan perhitungan sebagai berikut :
   $$Q3 = \frac{3}{4}N = \frac{3}{4}.40 = 20$$
   Jadi, kuartil ketiga (Q3) berada pada data ke-30.
2. Tentukan kelas Q3
   Untuk mengetahui kelas Q3, maka dilakukan penjumlahan frekuensi sebelum kelas kuartil , sebagai berikut :

| Kelas Interval | Frekuensi (f) | Frekuensi Kumulatif (F) |
|---|---|---|
| 2-4 | 5 | 5 |
| 5-7 | 8 | 13 |
| 8-10 | 12 | 25 |
| 11-13 | 10 | 35 |
| 14-16 | 5 | 40 |

Data ke-30 (Q3) berada di kelas 11-13, karena frekuensi kumulatif sebelumnya $(25) < 30 \leq F\ saat\ ini\ (35)$.

3. Gunakan rumus kuartil data kelompok

$$Qk = L + \left( \frac{\frac{kN}{4} - F}{f} \right) c$$

Untuk Q3 ($k = 3$) :

- L = 10,5 (batas bawah kelas Q3 adalah 11-0,5)
- N = 40
- F = 25 (frekuensi kumulatif sebelum kelas Q3)
- f = 10 (frekuensi kelas Q3)
- c = 2 (panjang kelas)

maka :

$$Q3 = 10{,}5 + \left( \frac{\frac{3.40}{4} - 25}{10} \right) . 2$$

$$Q2 = 10{,}5 + \left(\frac{\frac{120}{4} - 25}{10}\right).2$$

$$Q2 = 10{,}5 + \left(\frac{30 - 25}{10}\right).2$$

$$Q2 = 10{,}5 + \left(\frac{5}{10}\right).2$$

$$Q2 = 10{,}5 + \left(\frac{10}{10}\right) = 10{,}5 + 1 = 11{,}5$$

Dari perhitungan di atas, diperoleh Q3 dari pendapatan bulanan karyawan tersebut adalah Rp 11.500.000,-.

### 5.3.2 Rumus Desil (D)

**a. Rumus Desil untuk Data Tunggal**

Jika data berjumlah **n**, maka Desil dihitung dengan rumus (Walpole, 1992):

$$Dk = \frac{k(n+1)}{10}$$

Di mana:
$D_k$ = Nilai Desil ke-k
$k$ = Nomor Desil yang dicari (1 hingga 9),
$n$ = Jumlah data dalam kumpulan data.

Jika nilai yang dicari bukan bilangan bulat, maka dilakukan interpolasi untuk mendapatkan nilai yang lebih akurat (Sugiyono, 2007).

**Contoh Perhitungan 1**:
Seorang analis pemasaran ingin mengetahui distribusi pendapatan bulanan dari 10 pelaku UMKM di suatu daerah. Data pendapatan (dalam juta rupiah) adalah sebagai berikut:

5, 3, 5, 2, 6, 7, 8, 15, 12, 10

Tentukan Desil ke-7 dari data tersebut .

**Langkah-langkah Menghitung Desil:**

**Langkah 1**: Urutkan data dari yang terkecil ke terbesar

2, 3, 5, 5, 6, 7, 8, 10, 12, 15

**Langkah 2**: Gunakan rumus desil data tunggal:

$$Dk = \frac{k(n+1)}{10}$$

Dk: D7
K : 7
N : 10

maka :

$$D7 = \frac{7(10+1)}{10}$$

$$D7 = \frac{7(11)}{10} = \frac{77}{10} = 7{,}7$$

**Langkah 3**: Cari nilai D7 di posisi ke-7,7

Hasil 7,7 menunjukkan bahwa posisi Desil ke-7 (D7) terletak di antara data ke-7 dan ke-8, yaitu 8 (X7) dan 10 (X8).

2, 3, 5, 5, 6, 7, 8, 10, 12, 15

**Gunakan interpolasi**:

$$D7 = X7 + 0{,}7(X8 - X7)$$
$$(Catatan: 0{,}7\ berasal\ dari\ pemisahan\ 7{,}7\ (7 + 0{,}7)$$
$$D7 = 8 + 0{,}7(10 - 8) = 8 + 0{,}7(2)$$
$$D7 = 8 + 1{,}4 = 9{,}4$$

Maka Desil ke-7 (D7) terletak di posisi 9,4.

Jadi Desil ke-7 (D7) dari distribusi pendapatan bulanan UMKM adalah Rp 9.400.000,-. Artinya, 70% pelaku UMKM memiliki pendapatan di bawah Rp 9.400.000,- dan sisanya di atas angka tersebut.

**Contoh Perhitungan 2**:
Sebuah penelitian dilakukan terhadap 15 usaha mikro untuk mengetahui distribusi pendapatan bulanan mereka (dalam juta rupiah). Tujuan penelitian adalah mengetahui Desil ke-4 (D4), yaitu pendapatan yang membagi 40% usaha dengan pendapatan di bawah nilai tersebut.

**Data Pendapatan Bulanan (juta rupiah)**
(sudah diurutkan dari yang terkecil ke terbesar):

1.2, 1.5, 1.6, 1.8, 2.0, 2.1, 2.2, 2.5, 2.7, 2.9, 3.0, 3.2, 3.5, 3.8, 4.0

Jumlah data n=15

**Langkah 1:** Gunakan Rumus Desil
Rumus posisi desil ke-k:

$$Dk = \frac{k(n+1)}{10}$$

Untuk Desil ke-4 (D4):

$$D4 = \frac{4(15+1)}{10}$$

$$D4 = \frac{4(16)}{10} = \frac{64}{10} = 6{,}4$$

**Langkah 2**: Cari nilai D4 di posisi ke-6,4

Hasil 6,4 menunjukkan bahwa posisi Desil ke-4 (D4) terletak di antara data ke-6 dan ke-7, yaitu 2,1 (X6) dan 2,2 (X7).

1.2, 1.5, 1.6, 1.8, 2.0, 2.1, 2.2, 2.5, 2.7, 2.9, 3.0, 3.2, 3.5, 3.8, 4.0

**Gunakan interpolasi**:

$$D4 = X6 + 0{,}4(X7 - X6)$$

$(Catatan: 0{,}4\ berasal\ dari\ pemisahan\ 6{,}4\ (6 + 0{,}4)$

$$D4 = 2{,}1 + 0{,}4(2{,}2 - 2{,}1) = 2{,}1 + 0{,}4(0{,}1)$$

$$D4 = 2{,}1 + 0{,}04 = 2{,}14$$

Maka Desil ke-4 (D4) terletak di posisi 2,14.

Jadi Desil ke-4 (D4) dari distribusi pendapatan bulanan UMKM adalah Rp 2.140.000,-. Artinya, 40% pelaku usaha mikro memiliki pendapatan di bawah Rp 2.140.000,- dan sisanya di atas angka tersebut.

**b. Rumus Desil untuk Data Kelompok**

Jika data berbentuk distribusi frekuensi, kuartil dihitung menggunakan rumus (Triola, 2014):

$$Dk = L + \left(\frac{\frac{kN}{10} - F}{f}\right)c$$

dengan:
*Dk* = Desil *ke-k* (dengan *k=1,2, …., 9*).
*L* = Batas bawah kelas desil.
*N* = Jumlah total frekuensi (jumlah seluruh data).
*F* = Frekuensi kumulatif sebelum kelas desil.
*f*= Frekuensi pada kelas desil.
*c* = Panjang interval kelas.
*k* = Urutan desil yang dicari (1 hingga 9).

**Contoh Perhitungan** :
Seorang manajer HRD perusahaan ingin mengetahui distribusi pendapatan karyawan dalam kelompok-kelompok desil (D1, D5, dan D9) agar dapat menyusun strategi pemberian insentif yang adil. Berikut adalah data pendapatan bulanan karyawan (dalam juta rupiah):

| Kelas Interval | Frekuensi (f) |
|---|---|
| 2 - 4 | 6 |
| 5 - 7 | 10 |
| 8 - 10 | 14 |
| 11 - 13 | 8 |
| 14 - 16 | 4 |

Jumlah frekuensi (N) = 6+10+14+8+4=42

**Menghitung Desil ke-1 (D1)**

Langkah-langkah Perhitungan:

1. Tentukan posisi desil
   Untuk mengetahui letak posisi Desil 1 (D1), maka dilakukan perhitungan sebagai berikut :

$$Dk = \frac{k}{10} . N$$

$$D1 = \frac{1}{10} .42 = 4{,}2$$

   Jadi, Desil ke-1 (D1) berada pada data ke-4,2.

2. Tentukan kelas D1
   Untuk mengetahui kelas D1, maka dilakukan penjumlahan frekuensi sebelum kelas desil , sebagai berikut :

| Kelas Interval | Frekuensi (f) | Frekuensi Kumulatif (F) |
|---|---|---|
| 2 - 4 | 6 | 6 |
| 5 - 7 | 10 | 16 |
| 8 - 10 | 14 | 30 |
| 11 - 13 | 8 | 38 |
| 14 - 16 | 4 | 42 |

   Data ke-4,2 (D1) berada di kelas 2-4, karena frekuensi kumulatif sebelumnya $(0) < 4{,}2 \leq F\ saat\ ini\ (6)$.

3. Gunakan rumus desil data kelompok

$$Dk = L + \left(\frac{\frac{kN}{10} - F}{f}\right)c$$

dengan:
$Dk$ = D1
$L$ = 1,5 (batas bawah kelas D1 adalah 2-0,5)
$N$ = 42
$F$ = 0 (frekuensi kumulatif sebelum kelas D1)
$f$ = 6
$c$ = 2
$k$ = 1
maka :

$$D1 = 1{,}5 + \left(\frac{\frac{1.42}{10} - 0}{6}\right)2$$

$$D1 = 1{,}5 + \left(\frac{\frac{42}{10}}{6}\right)2$$

$$D1 = 1{,}5 + \left(\frac{42.2}{60}\right) = 1{,}5 + \left(\frac{84}{60}\right) = 1{,}5 + 1{,}4 = 2{,}9$$

Berdasarkan perhitungan di atas, diperoleh distribusi pendapatan karyawan dalam kelompok desil ke-1 (D1) adalah Rp 2.900.000,-. Artinya, 10% karyawan memiliki pendapatan kurang dari Rp 2.900.000,- dan sisanya di atas angka tersebut.

**Menghitung Desil ke-5 (D5)**

Langkah-langkah Perhitungan:

1. Tentukan posisi desil
   Untuk mengetahui letak posisi Desil 5 (D5), maka dilakukan perhitungan sebagai berikut :

$$Dk = \frac{k}{10} . N$$

$$D5 = \frac{5}{10} .42 = 21$$

   Jadi, Desil ke-5 (D5) berada pada data ke-21.

2. Tentukan kelas D5
   Untuk mengetahui kelas D5, maka dilakukan penjumlahan frekuensi sebelum kelas desil , sebagai berikut :

| Kelas Interval | Frekuensi (f) | Frekuensi Kumulatif (F) |
|---|---|---|
| 2 - 4 | 6 | 6 |
| 5 - 7 | 10 | 16 |
| 8 - 10 | 14 | 30 |
| 11 - 13 | 8 | 38 |
| 14 - 16 | 4 | 42 |

   Data ke-21 (D5) berada di kelas 8-10, karena frekuensi kumulatif sebelumnya $(16) < 21 \leq F\ saat\ ini\ (30)$.

3. Gunakan rumus desil data kelompok

$$Dk = L + \left(\frac{\frac{kN}{10} - F}{f}\right)c$$

dengan:
$Dk$ = D5
$L$ = 7,5 (batas bawah kelas D5 adalah 8-0,5)
$N$ = 42
$F$ = 16 (frekuensi kumulatif sebelum kelas D5)
$f$ = 14
$c$ = 2
$k$ = 5
maka :

$$D5 = 7{,}5 + \left(\frac{\frac{5.42}{10} - 16}{14}\right)2$$

$$D5 = 7{,}5 + \left(\frac{\frac{210}{10} - 16}{14}\right)2 = 7{,}5 + \left(\frac{21 - 16}{14}\right)2$$

$$D5 = 7{,}5 + \left(\frac{5}{14}\right)2 = 7{,}5 + \left(\frac{10}{14}\right) = 7{,}5 + 0{,}714 = 8{,}214$$

Berdasarkan perhitungan di atas, diperoleh distribusi pendapatan karyawan dalam kelompok desil ke-5 (D5) adalah Rp 8.214.000,-. Artinya, 50% karyawan memiliki pendapatan kurang dari Rp 8.214.000,- dan sisanya di atas angka tersebut.

**Menghitung Desil ke-9 (D9)**

Langkah-langkah Perhitungan:

1. Tentukan posisi desil
   Untuk mengetahui letak posisi Desil 9 (D9), maka dilakukan perhitungan sebagai berikut :

$$Dk = \frac{k}{10} . N$$

$$D9 = \frac{9}{10} .42 = 37{,}8$$

   Jadi, Desil ke-9 (D9) berada pada data ke-37,8.

2. Tentukan kelas D9
   Untuk mengetahui kelas D9, maka dilakukan penjumlahan frekuensi sebelum kelas desil , sebagai berikut :

| Kelas Interval | Frekuensi (f) | Frekuensi Kumulatif (F) |
|---|---|---|
| 2 - 4 | 6 | 6 |
| 5 - 7 | 10 | 16 |
| 8 - 10 | 14 | 30 |
| 11 - 13 | 8 | 38 |
| 14 - 16 | 4 | 42 |

   Data ke-37,8 (D9) berada di kelas 11-13, karena frekuensi kumulatif sebelumnya $(30) < 37{,}8 \leq F\ saat\ ini\ (38)$.

3. Gunakan rumus desil data kelompok

$$Dk = L + \left(\frac{\frac{kN}{10} - F}{f}\right)c$$

dengan:

$Dk$ = D9
$L$ = 10,5 (batas bawah kelas D9 adalah 11-0,5)
$N$ = 42
$F$ = 30 (frekuensi kumulatif sebelum kelas D9)
$f$ = 8
$c$ = 2
$k$ = 9
maka :

$$D9 = 10{,}5 + \left(\frac{\frac{9.42}{10} - 30}{8}\right)2$$

$$D9 = 10{,}5 + \left(\frac{\frac{378}{10} - 30}{8}\right)2 = 10{,}5 + \left(\frac{37{,}8 - 30}{8}\right)2$$

$$D9 = 10{,}5 + \left(\frac{7{,}8}{8}\right)2 = 10{,}5 + \left(\frac{15{,}6}{8}\right) = 10{,}5 + 1{,}95 = 12{,}45$$

Berdasarkan perhitungan di atas, diperoleh distribusi pendapatan karyawan dalam kelompok desil ke-9 (D9) adalah Rp 12.450.000,-. Artinya, 90% karyawan memiliki pendapatan kurang dari Rp 12.450.000,- dan sisanya di atas angka tersebut.

### 5.3.3 Rumus Persentil (P)

**a. Rumus Persentil untuk Data Tunggal**

Untuk menghitung persentil ke-k dari suatu kumpulan data berukuran n, digunakan rumus (Walpole, 1992):

$$Pk = \frac{k(n+1)}{100}$$

Di mana:
$P_k$ = Nilai Persentil ke-k
$k$ = Nomor Persentil yang dicari (1 hingga 99),
$n$ = Jumlah data dalam kumpulan data.
Jika hasil perhitungan bukan bilangan bulat, interpolasi digunakan untuk mendapatkan nilai yang lebih akurat (Sugiyono, 2007).

**Contoh Perhitungan 1**:

Seorang analis ekonomi ingin mengetahui P75 (persentil ke-75) dari pendapatan tahunan (dalam juta rupiah) dari 10 individu berikut ini :

18, 15, 12, 22, 25, 38, 33, 30, 50, 40

**Langkah Perhitungan**

1. Urutkan data pendapatan tersebut menjadi :

12, 15, 18, 22, 25, 30, 33, 38, 40, 50

2. Gunakan rumus Persentil data Tunggal

$$Pk = \frac{k(n+1)}{100}$$

Di mana:
$P_{75}$= Nilai Persentil ke-75
$k$ = 75
$n$ = 10
maka :

$$P75 = \frac{75(10+1)}{100}$$

$$P75 = \frac{75(11)}{100} = \frac{825}{100} = 8{,}25$$

Posisi 8,25 berarti nilai persentil terletak di antara data ke-8 dan ke-9, yaitu 38 (X8) dan 40 (X9).

12, 15, 18, 22, 25, 30, 33, 38, 40, 50

3. Lakukan Interpolasi

$$P75 = X8 + (0{,}25)(X9 - X8)$$

$(Catatan: 0{,}25\ berasal\ dari\ pemisahan\ 8{,}25\ (8 + 0{,}25)$

$$P75 = 38 + (0{,}25)(40 - 38) = 38 + (0{,}25)(2)$$

$$P75 = 38 + 0{,}5 = 38{,}5$$

Persentil ke-75 menunjukkan bahwa 75% dari individu memiliki pendapatan kurang dari atau sama dengan Rp 38.500.000,-, sedangkan 25% sisanya memiliki pendapatan lebih besar dari angka ini.

**Contoh Perhitungan 2**:
Sebuah perusahaan ingin menetapkan harga jual produk berdasarkan distribusi harga pasar (dalam ribuan rupiah) dari 15 kompetitor berikut :

78, 55, 110, 62, 65, 75, 72, 70, 50, 90, 85, 80, 95, 100, 60

Karena Perusahaan ingin mengetahui batas harga premium harga jual produk, maka tentukanlah P90 (persentil ke-90) berdasarkan data harga pasar tersebut.

**Langkah Perhitungan**

1. Urutkan data harga pasar tersebut menjadi :

   50, 55, 60, 62, 65, 70, 72, 75, 78, 80, 85, 90, 95, 100, 110

2. Gunakan rumus Persentil data Tunggal

$$Pk = \frac{k(n+1)}{100}$$

Di mana:
$P_{90}$= Nilai Persentil ke-90
$k$ = 90
$n$ = 15
maka :

$$P90 = \frac{90(15+1)}{100}$$

$$P90 = \frac{90(16)}{100} = \frac{1440}{100} = 14{,}4$$

Posisi 14,4 berarti nilai persentil terletak di antara data ke-14 dan ke-15, yaitu 100 (X14) dan 110 (X15).

50, 55, 60, 62, 65, 70, 72, 75, 78, 80, 85, 90, 95, 100, 110

3. Lakukan Interpolasi

$$P90 = X14 + (0{,}4)(X15 - X14)$$

$(Catatan : 0{,}4\ berasal\ dari\ pemisahan\ 14{,}4\ (14 + 0{,}4)$

$$P90 = 100 + (0{,}4)(110 - 100) = 100 + (0{,}4)(10)$$

$$P90 = 100 + 4 = 104$$

Persentil ke-90 menunjukkan bahwa 90% dari harga produk kompetitor berada di bawah Rp 104.000,-, sehingga perusahaan dapat menetapkan harga premium di atas nilai ini.

**b. Rumus Persentil untuk Data Kelompok**

Jika data berbentuk distribusi frekuensi, maka persentil dihitung menggunakan rumus:

$$Pk = L + \left( \frac{\frac{kN}{100} - F}{f} \right) x\ c$$

dengan:

*Pk* = Persentil *ke-k* (dengan *k=1,2, …., 99*).
*L* = Batas bawah kelas persentil.
*N* = Jumlah total frekuensi (jumlah seluruh data).
*F* = Frekuensi kumulatif sebelum kelas persentil.
*f*= Frekuensi pada kelas persentil.
*c* = Panjang interval kelas.
*k* = Urutan persentil yang dicari (1 hingga 99).

**Contoh Perhitungan** :

Seorang ekonom ingin mengevaluasi distribusi pengeluaran konsumsi rumah tangga per bulan (dalam juta rupiah) untuk memahami tingkat kesejahteraan berdasarkan kelompok persentil (P25, P50, dan P90). Data hasil survei dari 50 rumah tangga dirangkum sebagai berikut:

| Kelas Interval | Frekuensi (f) |
|---|---|
| 1 - 3 | 5 |
| 4 - 6 | 9 |
| 7 - 9 | 15 |
| 10 - 12 | 13 |
| 13 - 15 | 8 |

Jumlah frekuensi (N) = 5+9+15+13+8=50

**Menghitung Persentil ke-25 (P25)**

Langkah-langkah Perhitungan:

1. Tentukan posisi Persentil

Untuk mengetahui letak posisi Persentil ke-25 (P25), maka dilakukan perhitungan sebagai berikut :

$$Pk = \frac{k}{100} . N$$

$$P25 = \frac{25}{100} .50 = 12{,}5$$

Jadi, Persentil ke-25 (P25) berada pada data ke-12,5.

2. Tentukan kelas P25

Untuk mengetahui kelas P25, maka dilakukan penj $i$ (14).

3. Gunakan rumus Persentil data kelompok

$$Pk = L + \left( \frac{\frac{kN}{100} - F}{f} \right) c$$

dengan:

$Pk$ = P25

$L$ = 3,5 (batas bawah kelas P25 adalah 4-0,5)

$N$ = 50

$F$ = 5 (frekuensi kumulatif sebelum kelas P25)

$f$ = 9

$c$ = 3

$k$ = 25

maka :

$$P25 = 3{,}5 + \left( \frac{\frac{25.50}{100} - 5}{9} \right) 3$$

$$P25 = 3{,}5 + \left(\frac{\frac{1250}{100} - 5}{9}\right)3 = 3{,}5 + \left(\frac{12{,}5 - 5}{9}\right)3$$

$$P25 = 3{,}5 + \left(\frac{7{,}5}{9}\right)3 = 3{,}5 + \left(\frac{22{,}5}{9}\right) = 3{,}5 + 2{,}5 = 6$$

Berdasarkan perhitungan di atas, diperoleh distribusi pengeluaran konsumsi rumah tangga per bulan untuk memahami tingkat kesejahteraan berdasarkan kelompok persentil ke-25 (P25) adalah Rp 6.000.000,-. Artinya, 25% distribusi pengeluaran konsumsi rumah tangga per bulan kurang dari Rp 6.000.000,- dan sisanya di atas angka tersebut.

**Menghitung Persentil ke-50 (P50)**

Langkah-langkah Perhitungan:

1. Tentukan posisi Persentil

Untuk mengetahui letak posisi Persentil ke-50 (P50), maka dilakukan perhitungan sebagai berikut :

$$Pk = \frac{k}{100} \,.N$$

$$P50 = \frac{50}{100} \,.50 = 25$$

Jadi, Persentil ke-50 (P50) berada pada data ke-25.

2. Tentukan kelas P50

Untuk mengetahui kelas P50, maka dilakukan penjumlahan frekuensi sebelum kelas persentil , sebagai berikut :

| Kelas Interval | Frekuensi (f) | Frekuensi Kumulatif (F) |
|---|---|---|
| 1 - 3 | 5 | 5 |
| 4 - 6 | 9 | 14 |
| 7 - 9 | 15 | 29 |
| 10 - 12 | 13 | 42 |
| 13 - 15 | 8 | 50 |

Data ke-25 (P50) berada di kelas 7-9, karena frekuensi kumulatif sebelumnya $(14) < 25 \leq F\ saat\ ini\ (29)$.

3. Gunakan rumus Persentil data kelompok

$$Pk = L + \left(\frac{\frac{kN}{100} - F}{f}\right)c$$

dengan:
$Pk$ = P50
$L$ = 6,5 (batas bawah kelas P50 adalah 7-0,5)
$N$ = 50
$F$ = 14 (frekuensi kumulatif sebelum kelas P50)
$f$ = 15
$c$ = 3
$k$ = 50
maka :

$$P50 = 6{,}5 + \left(\frac{\frac{50.50}{100} - 5}{15}\right)3$$

$$P50 = 6{,}5 + \left(\frac{\frac{2500}{100} - 14}{15}\right)3 = 6{,}5 + \left(\frac{25-14}{15}\right)3$$

$$P50 = 6{,}5 + \left(\frac{11}{15}\right)3 = 6{,}5 + \left(\frac{33}{15}\right) = 6{,}5 + 2{,}2 = 8{,}7$$

Berdasarkan perhitungan di atas, diperoleh distribusi pengeluaran konsumsi rumah tangga per bulan untuk memahami tingkat kesejahteraan berdasarkan kelompok persentil ke-50 (P50) adalah Rp 8.700.000,-. Artinya, 50% distribusi pengeluaran konsumsi rumah tangga per bulan kurang dari Rp 8.700.000,- dan sisanya di atas angka tersebut.

**Menghitung Persentil ke-90 (P90)**

Langkah-langkah Perhitungan:

1. Tentukan posisi Persentil

Untuk mengetahui letak posisi Persentil ke-90 (P90), maka dilakukan perhitungan sebagai berikut :

$$Pk = \frac{k}{100} \, . N$$

$$P90 = \frac{90}{100} \, .50 = 45$$

Jadi, Persentil ke-90 (P90) berada pada data ke-45.

2. Tentukan kelas P90

Untuk mengetahui kelas P90, maka dilakukan penjumlahan frekuensi sebelum kelas persentil , sebagai berikut :

| Kelas Interval | Frekuensi (f) | Frekuensi Kumulatif (F) |
|---|---|---|
| 1 - 3 | 5 | 5 |
| 4 - 6 | 9 | 14 |
| 7 - 9 | 15 | 29 |
| 10 - 12 | 13 | 42 |
| 13 - 15 | 8 | 50 |

Data ke-45 (P90) berada di kelas 13-15, karena frekuensi kumulatif sebelumnya $(42) < 45 \leq F\ saat\ ini\ (50)$.

3. Gunakan rumus Persentil data kelompok

$$Pk = L + \left(\frac{\frac{kN}{100} - F}{f}\right)c$$

dengan:
$Pk$ = P90
$L$ = 12,5 (batas bawah kelas P90 adalah 13-0,5)
$N$ = 50
$F$ = 42 (frekuensi kumulatif sebelum kelas P90)
$f$ = 8
$c$ = 3
$k$ = 90

maka :

$$P90 = 12{,}5 + \left(\frac{\frac{90.50}{100} - 42}{8}\right)3$$

$$P90 = 12{,}5 + \left(\frac{\frac{4500}{100} - 42}{8}\right)3 = 12{,}5 + \left(\frac{45-42}{8}\right)3$$

$$P90 = 12{,}5 + \left(\frac{3}{8}\right)3 = 12{,}5 + \left(\frac{9}{8}\right) = 12{,}5 + 1{,}125 = 13{,}625$$

Berdasarkan perhitungan di atas, diperoleh distribusi pengeluaran konsumsi rumah tangga per bulan untuk memahami tingkat kesejahteraan berdasarkan kelompok persentil ke-90 (P90) adalah Rp 13.625.000,-. Artinya, 90% distribusi pengeluaran konsumsi rumah tangga per bulan kurang dari Rp 13.625.000,- dan sisanya di atas angka tersebut.

## DAFTAR PUSTAKA

Sudjana (1996) *Metoda Statistika*. ke-6. Bandung: Tarsito.

Sugiyono (2007) *Statistika untuk Penelitian*. Bandung: Alfabeta.

Trihendradi, C. (2019) *Statistika Deskriptif dan Inferensial*. Yogyakarta: Penerbit Andi.

Triola, M.F. (2014) *Elementary Statistics*. Twelfth. Harlow: Pearson Education Limited.

Walpole, R.E. (1992) *Pengantar Statistika*. ke-3. Jakarta: PT Gramedia Pustaka Utama.

# BAB 6
# KONSEP DASAR PROBABILITAS

**Oleh Roosganda Elizabeth**

## 6.1 Pengantar Probabilitas

Probabilitas merupakan salah satu konsep dasar yang sangat penting dalam statistik, terutama dalam konteks pengambilan keputusan ekonomi dan bisnis. Ketika individu atau organisasi harus membuat keputusan berdasarkan informasi yang tidak lengkap atau tidak pasti, maka probabilitas menjadi alat bantu utama untuk menilai kemungkinan berbagai hasil. Dalam lingkungan bisnis modern yang serba cepat dan kompleks, pemahaman yang kuat terhadap konsep probabilitas memberikan keunggulan kompetitif yang signifikan. Dalam konteks ekonomi dan bisnis, pemahaman tentang probabilitas sangat penting untuk pengambilan keputusan di bawah ketidakpastian. Mulai dari peramalan permintaan, evaluasi risiko investasi, hingga penentuan strategi pemasaran, semua melibatkan elemen probabilitas.

Dalam kehidupan sehari-hari, probabilitas digunakan untuk memperkirakan kemungkinan suatu kejadian. Misalnya, seorang investor mempertimbangkan peluang keuntungan dan kerugian dari investasi di pasar saham. Seorang pemilik usaha restoran mencoba memperkirakan jumlah pelanggan yang akan datang dalam sehari. Sebagai

contoh, seorang manajer pemasaran ingin mengetahui peluang keberhasilan dari peluncuran produk baru. Dengan menggunakan pendekatan probabilistik berdasarkan data historis dan tren pasar, manajer tersebut dapat membuat keputusan yang lebih terinformasi.

Dalam statistik ekonomi dan bisnis, probabilitas digunakan untuk mendukung pengambilan keputusan dalam berbagai situasi, seperti:

1. Evaluasi Risiko: Dalam manajemen risiko, perusahaan menggunakan probabilitas untuk mengukur potensi kerugian atau kegagalan.
2. Peramalan dan Prediksi: Perusahaan menggunakan model probabilistik untuk memperkirakan permintaan produk, tren pasar, dan perubahan harga.
3. Pengendalian Kualitas: Dalam proses produksi, probabilitas digunakan untuk menentukan apakah suatu produk cacat berada dalam batas yang dapat diterima.
4. Strategi Pemasaran: Probabilitas membantu dalam menilai keberhasilan kampanye pemasaran berdasarkan data historis pelanggan dan tren perilaku konsumen.

Untuk memahami probabilitas secara komprehensif, diperlukan pemahaman tentang berbagai istilah dasar seperti eksperimen, hasil, ruang sampel, dan kejadian. Bab-bab selanjutnya akan mengupas lebih dalam mengenai konsep-konsep tersebut beserta penerapannya dalam konteks ekonomi dan bisnis.

Selain itu, penting juga untuk menyadari bahwa pendekatan probabilitas tidak selalu bersifat deterministik, tetapi lebih bersifat inferensial. Artinya, dalam banyak kasus, kita membuat estimasi atau penilaian terhadap probabilitas berdasarkan sampel data atau pengalaman masa lalu. Oleh karena itu, penguasaan probabilitas membuka jalan bagi pemahaman yang lebih dalam tentang statistik inferensial,

regresi, dan analisis prediktif yang sangat dibutuhkan dalam dunia bisnis kontemporer.

Kesimpulannya, probabilitas adalah dasar penting dalam pemodelan ketidakpastian. Dalam dunia yang penuh dengan variabilitas, ketidakpastian, dan kompleksitas, pemahaman terhadap probabilitas tidak hanya memberikan pemahaman konseptual, tetapi juga keterampilan praktis yang dapat diterapkan untuk mengambil keputusan yang lebih baik di dunia nyata.

## 6.2 Konsep Dasar dalam Probabilitas

Agar dapat memahami probabilitas secara menyeluruh, penting untuk mengenali unsur-unsur dasarnya yang menjadi fondasi dalam setiap perhitungan peluang. Konsep dasar ini mencakup eksperimen, hasil, ruang sampel, dan kejadian. Masing-masing komponen ini akan dijelaskan secara rinci berikut ini.

a. Eksperimen, Hasil, dan Ruang Sampel

- Dalam statistik, eksperimen merupakan suatu proses yang menghasilkan satu atau lebih hasil, tidak harus bersifat ilmiah; Eksperimen bisa berupa aktivitas sederhana seperti melempar koin atau survei terhadap preferensi konsumen.
- Hasil (*outcome*) adalah hasil spesifik dari suatu eksperimen, yang merupakan keluaran spesifik dari suatu eksperimen. Sebagai contoh, hasil dari pelemparan satu koin bisa berupa "muka" atau "angka".

Ruang sampel biasanya dilambangkan dengan huruf kapital S. Contoh: ruang sampelya pelemparan satu dadu adalah

$$S = \{1, 2, 3, 4, 5, 6\}$$

Penting untuk memahami bahwa ruang sampel mencakup keseluruhan kemungkinan, dan setiap kejadian yang dihitung probabilitasnya merupakan bagian dari ruang sampel ini.

b. Kejadian (*Event*)

Kejadian adalah himpunan bagian dari satu atau beberapa hasil dari ruang sampel. Kejadian dapat bersifat sederhana (satu hasil) atau majemuk (lebih dari satu hasil).
Contoh:
Kejadian A: Mendapat angka genap saat melempar dadu.
A = {2, 4, 6}

Kejadian dapat diklasifikasikan menjadi:

- Kejadian Sederhana (*Simple Event*): Kejadian yang hanya terdiri atas satu hasil tunggal.
- Kejadian Majemuk (*Compound Event*): Kejadian yang terdiri atas dua atau lebih hasil.

Kejadian juga dapat diklasifikasikan berdasarkan keterkaitannya dengan kejadian lain, seperti:

- Kejadian Komplementer: Jika A adalah suatu kejadian, maka komplementernya adalah kejadian yang mencakup semua hasil yang tidak termasuk dalam A. Komplementer A ditulis sebagai A'.

c. Peluang (Probabilitas)

Probabilitas suatu kejadian A, ditulis P(A), adalah ukuran kuantitatif dari kemungkinan terjadinya A. Dalam pendekatan klasik, Jika semua hasil dalam ruang sampel memiliki peluang yang sama, maka: probabilitas suatu kejadian dapat dihitung dengan:
P(A) = (Jumlah hasil dalam A) / (Jumlah total hasil dalam S)

Contoh: Dalam pelemparan sebuah dadu, probabilitas muncul angka 2 adalah: P(2) = 1/6
Untuk kejadian majemuk seperti mendapatkan angka genap,

d. Notasi dan Representasi Probabilitas

Contoh P(A) = 0,75 atau 75%. Notasi P(A ∩ B) menyatakan probabilitas kejadian A dan B terjadi secara bersamaan, sementara P(A ∪ B) menyatakan probabilitas A atau B terjadi.
Pemahaman terhadap simbol-simbol ini sangat penting karena dalam penerapan lanjut seperti probabilitas bersyarat dan distribusi probabilitas, notasi tersebut akan digunakan secara konsisten.

e. Aplikasi dalam Dunia Nyata

Dalam konteks ekonomi dan bisnis, konsep dasar probabilitas ini digunakan dalam berbagai situasi:

- Perusahaan asuransi menentukan premi berdasarkan probabilitas terjadinya klaim.
- Perusahaan logistik memprediksi kemungkinan keterlambatan pengiriman.
- Lembaga keuangan mengevaluasi kemungkinan gagal bayar dari kreditur.

Konsep dasar probabilitas membentuk fondasi penting dalam memahami analisis risiko dan model prediktif, serta menjadi dasar bagi metode kuantitatif lanjutan seperti distribusi peluang, analisis regresi, dan simulasi Monte Carlo. Dengan memahami elemen-elemen dasar ini secara menyeluruh, pembaca akan siap untuk mendalami aturan dan teknik lanjutan dalam probabilitas yang akan dibahas dalam bab-bab berikutnya.

## 6.3 Aturan Dasar Probabilitas

Dalam teori probabilitas, terdapat beberapa aturan dasar yang sangat penting dan digunakan secara luas dalam penghitungan peluang, baik dalam kasus sederhana maupun kompleks. Aturan-aturan ini membantu kita memahami bagaimana peluang dari kejadian tunggal maupun gabungan dihitung, serta bagaimana peluang satu kejadian dapat memengaruhi peluang kejadian lainnya.

a. Aturan Penjumlahan (*Addition Rule*)

Aturan penjumlahan digunakan untuk menentukan probabilitas dari kejadian gabungan, yaitu kejadian A atau B terjadi (A ∪ B). Aturan ini memiliki dua bentuk, tergantung pada apakah kejadian tersebut saling lepas (*mutually exclusive*) atau tidak.

1) Kejadian Saling Lepas (*Mutually Exclusive Events*)
Dua kejadian A dan B dikatakan saling lepas jika keduanya tidak dapat terjadi secara bersamaan. Dalam kasus ini:
$P(A \cup B) = P(A) + P(B)$
Probabilitas pelemparan sebuah dadu mendapatkan angka 2 (A) atau angka 5 (B):
$P(A \cup B) = P(2) + P(5) = 1/6 + 1/6 = 1/3$

2) Jika kejadian A dan B dapat terjadi secara bersamaan sebagai Kejadian Tidak Saling Lepas, maka:
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Contoh: Dalam sebuah survei, diketahui bahwa 60% responden menyukai kopi (A), 40% menyukai teh (B), dan 20% menyukai keduanya (A ∩ B). Maka, probabilitas seseorang menyukai kopi atau teh: $P(A \cup B) = 0,6 + 0,4 - 0,2 = 0,8$

b. Aturan Perkalian (*Multiplication Rule*)

Sama seperti penjumlahan, aturan ini juga tergantung apakah kejadian A dan B bersifat independen atau tidak.

1) Kejadian Independen Dua kejadian A dan B dikatakan independen jika kejadian satu tidak memengaruhi kejadian lainnya. Dalam hal ini:
$P(A \cap B) = P(A) \times P(B)$

Contoh: Probabilitas mendapatkan angka genap dalam pelemparan satu dadu ($P(A) = 3/6$) dan mendapatkan kepala dalam pelemparan koin ($P(B) = 0,5$): $P(A \cap B) = (3/6) \times (0,5) = 0,25$

2) Kejadian Tidak Independen (Bersyarat) Jika kejadian A dan B saling bergantung, maka digunakan probabilitas bersyarat:
$P(A \cap B) = P(A) \times P(B|A)$

di mana $P(B|A)$ adalah probabilitas B terjadi setelah A terjadi.

Contoh: Dalam suatu perusahaan, 70% karyawan mengikuti pelatihan (A), dan dari mereka, 60% berhasil mencapai target kerja (B). Maka probabilitas karyawan mengikuti pelatihan dan mencapai target kerja: $P(A \cap B) = 0,7 \times 0,6 = 0,42$

c. Aturan Probabilitas Komplementer

Probabilitas komplementer suatu kejadian dengan lambang A' :
$P(A') = 1 - P(A)$
Aturan ini sangat berguna ketika lebih mudah menghitung probabilitas kejadian tidak terjadi dibanding kejadian itu sendiri.

Contoh: Probabilitas seseorang gagal dalam wawancara kerja adalah 0,25. Maka, probabilitas orang tersebut berhasil dalam wawancara: $P(A') = 1 - 0{,}25 = 0{,}75$

d. Aplikasi Aturan dalam Studi Kasus Bisnis

1. Manajemen Risiko: Dalam manajemen risiko keuangan, kombinasi aturan perkalian dan penjumlahan digunakan untuk menilai peluang kerugian portofolio investasi yang terdiri dari beberapa instrumen berbeda.
2. Perencanaan Pemasaran: Misalnya, tim pemasaran ingin mengetahui probabilitas keberhasilan kampanye pada dua segmen pasar yang berbeda. Dengan mengetahui apakah segmen tersebut independen atau tidak, maka pendekatan perhitungan bisa disesuaikan.
3. Pengambilan Keputusan Operasional: Dalam produksi, probabilitas kegagalan komponen dihitung menggunakan aturan-aturan dasar probabilitas untuk memperkirakan risiko kegagalan sistem.

e. Pentingnya Pemahaman Terhadap Aturan Dasar

Pemahaman terhadap aturan dasar probabilitas menjadi fondasi dalam analisis statistik lanjutan, seperti teori distribusi, statistika inferensial, dan pengambilan keputusan berbasis data (*data-driven decision making*). Dalam praktiknya, kesalahan dalam menerapkan aturan probabilitas dapat mengarah pada interpretasi yang keliru dan keputusan bisnis yang tidak tepat.

Dengan penguasaan terhadap aturan dasar probabilitas ini, pelaku ekonomi dan bisnis dapat mengembangkan intuisi statistik yang kuat, sehingga mampu menilai risiko dan peluang secara rasional,

terukur, dan objektif dalam berbagai konteks pengambilan keputusan.

## 6.4 Probabilitas Bersyarat dan Ketergantungan

Dalam dunia nyata, banyak kejadian tidak terjadi secara independen. Misalnya, keberhasilan strategi pemasaran bisa tergantung pada musim, tren konsumen, atau kondisi ekonomi. Oleh karena itu, pemahaman tentang probabilitas bersyarat dan ketergantungan antar kejadian sangat penting dalam analisis statistik ekonomi dan bisnis.

a. Definisi Probabilitas Bersyarat

Rumus umum:

$P(A|B) = P(A \cap B) / P(B)$, dengan syarat $P(B) > 0$

Contoh: Dalam survei konsumen, diketahui bahwa 40% pelanggan membeli kopi, dan 25% membeli kopi serta makanan ringan. Maka, probabilitas seorang pelanggan membeli makanan ringan jika ia membeli kopi adalah:

$P(\text{Makanan Ringan}|\text{Kopi}) =$
$P(\text{Kopi} \cap \text{Makanan Ringan}) / P(\text{Kopi}) =$
$0{,}25 / 0{,}4 = 0{,}625$

Dimana dalam Probabilitas bersyarat (*conditional probability*), dengan asumsi kejadian B telah terjadi, maka peluang suatu kejadian A dengan Notasi P(A|B) = probabilitas A diberikan B.

b. Probabilitas Bersyarat dalam Konteks Bisnis

Probabilitas bersyarat digunakan secara luas dalam analisis bisnis, antara lain:

1. Analisis Konsumen: Memahami perilaku pembelian pelanggan berdasarkan atribut tertentu, misalnya usia, lokasi, atau waktu kunjungan.
2. Manajemen Persediaan: Menentukan permintaan

produk tertentu, jika diketahui permintaan produk lain meningkat.

3. Risiko Keuangan: Mengukur risiko gagal bayar kredit berdasarkan riwayat transaksi nasabah.

Contoh: Jika probabilitas seorang nasabah menunggak pinjaman adalah 0,1 dan dari mereka yang menunggak, 70% tidak memiliki jaminan, maka:

$$P(\text{Tidak Ada Jaminan}|\text{Menunggak}) = 0,7$$

c. Kejadian Independen dan Tidak Independen

Dua kejadian A dan B dikatakan independen jika kejadian A tidak memengaruhi probabilitas terjadinya kejadian B, dan sebaliknya. Dalam hal ini:

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

Bila kejadian saling tergantung (dependen), maka:

$$P(A \cap B) \neq P(A) \times P(B)$$

Contoh: Jika probabilitas seseorang membeli es krim adalah 0,3 dan membeli minuman dingin adalah 0,5, dan diketahui bahwa keduanya tidak saling memengaruhi, maka:

$$P(\text{Eskrim} \cap \text{Minuman Dingin}) = 0,3 \times 0,5 = 0,15$$

Namun, jika pembelian es krim cenderung terjadi bersamaan dengan minuman dingin (misalnya, $P(\text{Eskrim} \cap \text{Minuman Dingin}) = 0,25$), maka keduanya tidak independen.

d. Pengujian Ketergantungan dalam Data Ekonomi dan Bisnis

Untuk menguji apakah dua kejadian saling tergantung, kita bisa menggunakan perbandingan antara:

- $P(A \cap B)$
- $P(A) \times P(B)$

Jika nilainya sama, maka kejadian bersifat independen.

Contoh Praktis: Sebuah toko online mencatat bahwa:

- 60% pelanggan membeli produk A
- 40% membeli produk B
- 20% membeli keduanya

Maka:

- $P(A) = 0,6$
- $P(B) = 0,4$
- $P(A \cap B) = 0,2$
- $P(A) \times P(B) = 0,24$

Karena $0,2 \neq 0,24 \rightarrow$ A dan B tidak independen (terdapat ketergantungan antara dua kejadian).

e. Probabilitas Bersyarat dalam Pengambilan Keputusan Bisnis

Pemahaman terhadap probabilitas bersyarat membantu pengambil keputusan dalam:

1. Segmentasi Pasar: Menargetkan promosi berdasarkan kemungkinan pembelian produk lain.
2. Peluang Keberhasilan Proyek: Menghitung probabilitas keberhasilan suatu proyek berdasarkan keberhasilan tahap sebelumnya.
3. Audit dan Kontrol Risiko: Mengukur kemungkinan terjadinya kesalahan atau penipuan berdasarkan kejadian-kejadian sebelumnya.

f. Kesimpulan

Probabilitas bersyarat dan ketergantungan merupakan aspek penting dalam analisis data bisnis yang kompleks. Dalam era big data dan analitik prediktif, kemampuan mengidentifikasi hubungan antar variabel menjadi dasar dalam membangun model-model keputusan berbasis probabilitas yang akurat. Melalui penerapan konsep ini, pelaku ekonomi dan bisnis dapat membuat strategi yang lebih adaptif dan terukur terhadap dinamika pasar dan konsumen.

## 6.5 Pendekatan Probabilitas

Pendekatan probabilitas adalah metode untuk mengukur dan menafsirkan peluang terjadinya suatu kejadian. Dalam konteks statistik ekonomi dan bisnis, pendekatan ini penting karena membantu dalam pengambilan keputusan di bawah ketidakpastian. Tiga pendekatan utama digunakan dalam probabilitas: pendekatan klasik, pendekatan frekuensi relatif, dan pendekatan subjektif.

a. Pendekatan Klasik

Pendekatan klasik digunakan ketika semua kemungkinan hasil memiliki peluang yang sama (*equiprobable*). Pendekatan ini umumnya digunakan dalam permainan peluang atau situasi teoretis.

Rumus:

P(A) = Jumlah kejadian pendukung A / Jumlah total kejadian yang mungkin

Contoh: Jika sebuah dadu dilempar, maka peluang munculnya angka 4 adalah: P(4) = 1/6, karena terdapat 6 kemungkinan hasil yang setara.

Kelebihan:

- Mudah digunakan untuk peristiwa simetris.

Keterbatasan:

- Tidak relevan dalam kasus dunia nyata yang kompleks seperti perilaku konsumen atau tren pasar.

b. Pendekatan Frekuensi Relatif

Pendekatan ini mengukur probabilitas berdasarkan data historis atau eksperimen berulang. Semakin sering suatu kejadian terjadi dalam sejumlah besar percobaan, semakin besar probabilitasnya.

Rumus: P(A) = (Jumlah kejadian A terjadi) / (Jumlah

total percobaan)

Contoh: Jika dalam 1.000 transaksi, 150 transaksi merupakan pembelian produk A, maka probabilitas seorang pelanggan membeli produk A adalah: P(A) = 150/1.000 = 0,15

Kelebihan:

- Cocok untuk aplikasi bisnis dan ekonomi dengan data historis yang tersedia.

Keterbatasan:

- Tidak bisa digunakan jika tidak tersedia data atau jumlah percobaan sangat kecil.

c. Pendekatan Subjektif

Pendekatan berdasarkan penilaian pribadi/ intuisi/ keyakinan kemungkinan suatu peristiwa terjadi umum digunakan dalam keputusan manajerial ketika informasi kuantitatif terbatas.

Contoh: Seorang analis pasar memperkirakan bahwa kemungkinan keberhasilan peluncuran produk baru adalah 70%, berdasarkan pengalamannya, riset terbatas, dan wawancara konsumen.

Kelebihan:

- Berguna dalam situasi yang memerlukan keputusan cepat dan data tidak tersedia.

Keterbatasan:

- Rentan terhadap bias dan kurang objektif.

d. Perbandingan dan Aplikasi dalam Bisnis

| **Pendekatan** | **Cocok Untuk** | **Data Dibutuhkan** | **Risiko Bias** |
|---|---|---|---|
| Klasik | Kejadian teoritis | Tidak | Rendah |
| Frekuensi Relatif | Data historis dan eksperimen | Ya | Rendah |

| Pendekatan | Cocok Untuk | Data Dibutuhkan | Risiko Bias |
|---|---|---|---|
| Subjektif | Keputusan manajerial, intuisi | Tidak | Tinggi |

Dalam bisnis:

- Pendekatan klasik digunakan dalam model probabilitas dasar seperti asuransi.
- Pendekatan frekuensi relatif dipakai dalam analitik pemasaran dan manajemen risiko.
- Pendekatan subjektif dominan dalam keputusan investasi atau strategi saat data terbatas.

e. Kesimpulan

Pemilihan pendekatan probabilitas harus disesuaikan dengan konteks, data, dan kebutuhan analisis. Dalam praktik bisnis, sering kali digunakan kombinasi dari ketiga pendekatan untuk menghasilkan analisis yang komprehensif dan mendalam. Kemampuan memahami kelebihan dan keterbatasan masing-masing pendekatan akan meningkatkan kualitas pengambilan keputusan dalam menghadapi ketidakpastian di dunia ekonomi dan bisnis.

## 6.6 Teorema Bayes dan Aplikasinya dalam Bisnis

Teorema Bayes adalah salah satu konsep penting dalam teori probabilitas yang digunakan untuk memperbarui peluang suatu hipotesis berdasarkan informasi atau bukti baru. Dalam konteks bisnis dan ekonomi, teorema ini sangat berguna dalam pengambilan keputusan, terutama dalam kondisi ketidakpastian dan informasi yang berubah-ubah.

a. Konsep Teorema Bayes

Teorema Bayes memungkinkan kita menghitung probabilitas bersyarat suatu kejadian berdasarkan probabilitas awal (*prior probability*) dan informasi baru (*evidence*).

Rumus Teorema Bayes:

$$P(A|B) = / P(B)$$

Di mana:

- P(A|B) adalah probabilitas A terjadi jika B diketahui (probabilitas posterior).
- P(B|A) adalah probabilitas B terjadi jika A benar.
- P(A) adalah probabilitas awal dari A (probabilitas prior).
- P(B) adalah probabilitas total dari B.

b. Contoh Penerapan dalam Bisnis

Contoh 1: Analisis Risiko Kredit Bank ingin menilai apakah seorang peminjam berpotensi gagal bayar (default), berdasarkan data penghasilan dan riwayat kredit.

- P(Default) = 0,02 (2% dari pelanggan umumnya gagal bayar)
- P(Memiliki Riwayat Buruk | Default) = 0,90
- P(Memiliki Riwayat Buruk) = 0,05

Menggunakan Teorema Bayes:

P(Default | Riwayat Buruk) =

(0,90 × 0,02) / 0,05 = 0,018 / 0,05 = 0,36

Artinya, jika seorang calon nasabah memiliki riwayat kredit buruk, probabilitas dia gagal bayar naik dari 2% menjadi 36%.

Contoh 2: Strategi Pemasaran Sebuah perusahaan ingin memprediksi kemungkinan seorang pelanggan akan

membeli produk setelah melihat iklan daring.

- P(Membeli) = 0,10
- P(Melihat Iklan | Membeli) = 0,80
- P(Melihat Iklan) = 0,30

P(Membeli | Melihat Iklan) =

(0,80 × 0,10) / 0,30 = 0,08 / 0,30 = 0,267

Dengan kata lain, setelah pelanggan melihat iklan, peluang mereka untuk membeli meningkat dari 10% menjadi 26,7%.

c. Manfaat Teorema Bayes dalam Bisnis

1. Pengambilan Keputusan Lebih Tepat: Memberikan kerangka sistematis untuk memperbarui estimasi peluang dengan adanya data baru.
2. Manajemen Risiko: Membantu mengidentifikasi risiko berdasarkan gejala awal (indikator) seperti fraud detection atau default kredit.
3. Pemasaran Terarah: Mengarahkan promosi ke segmen pelanggan yang paling mungkin membeli berdasarkan respons masa lalu.
4. Peramalan: Digunakan dalam model prediktif dan *machine* learning untuk memperkirakan perilaku konsumen.

d. Keterbatasan Penggunaan Teorema Bayes

- Tergantung pada Probabilitas Awal: Jika *prior probability* tidak akurat, hasilnya juga bias.
- Kompleksitas Perhitungan: Untuk sistem besar dengan banyak variabel, perhitungan bisa sangat kompleks.
- Ketergantungan pada Data Historis: Bila data lama tidak relevan lagi, hasil perhitungan bisa menyesatkan.

e. Kesimpulan

Teorema Bayes merupakan alat penting dalam pengambilan keputusan berbasis data. Dengan memahami bagaimana informasi baru mempengaruhi estimasi peluang, pelaku bisnis dapat meningkatkan akurasi keputusan mereka. Meskipun memiliki keterbatasan, pendekatan bayesian tetap relevan dalam era big data dan analitik bisnis modern.

## 6.7 Distribusi Probabilitas Diskrit

Pendistribusian probabilitas di antara nilai-nilai diskrit dari sebuah variabel acak digambarkan sebagai Distribusi probabilitas diskrit, dimana variabel acak diskrit nya hanya dapat mengambil nilai-nilai tertentu dan terhitung, dalam satu kurun waktu.

a. Ciri-Ciri Distribusi Diskrit

1. Nilai variabel acak diskrit adalah nilai yang bisa dihitung (misalnya 0, 1, 2, 3,...).
2. Total probabilitas dari seluruh kemungkinan nilai adalah 1.
3. Setiap nilai individu dapat dihitung dan dijumlahkan probabilitasnya.

b. Fungsi Probabilitas (*Probability Mass Function* - PMF)

PMF menggambarkan probabilitas dari setiap nilai spesifik yang dapat diambil oleh variabel acak diskrit.
Contoh: Jika X adalah jumlah keluhan pelanggan dalam satu hari, maka fungsi probabilitas bisa berupa:

- P(X = 0) = 0,1
- P(X = 1) = 0,3
- P(X = 2) = 0,4
- P(X = 3) = 0,2

Jumlah seluruh probabilitas:

$$P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 1$$

c. Contoh Distribusi Diskrit dalam Bisnis

1. Distribusi Binomial: Digunakan ketika terdapat dua hasil kemungkinan (sukses/gagal). Contoh: probabilitas pelanggan membeli produk setelah mencoba sampel.

   Rumus:

   $$P(X = k) = C(n, k) * p^k * (1-p)^(n-k)$$

   Di mana:

   n = jumlah percobaan

   k = jumlah keberhasilan

   p = probabilitas keberhasilan dalam satu percobaan

2. Distribusi Poisson: Cocok untuk menghitung jumlah kejadian dalam suatu interval waktu atau ruang tertentu. Contoh: jumlah pelanggan datang ke restoran setiap jam.

   Rumus:

   $$P(X = x) = (e^(-\lambda) * \lambda^x) / x!$$

   Di mana λ adalah rata-rata kejadian per satuan waktu.

d. Aplikasi Distribusi Diskrit dalam Bisnis dan Ekonomi

- Manajemen Persediaan: Mengestimasi jumlah barang yang harus disediakan berdasarkan permintaan rata-rata (misal menggunakan distribusi Poisson).
- Kepuasan Pelanggan: Mengukur jumlah komplain pelanggan tiap minggu dan memprediksi kebutuhan peningkatan layanan.
- Keuangan: Menentukan probabilitas default pinjaman dalam portofolio kredit.

e. Analisis dan Interpretasi

Dengan menggunakan distribusi diskrit, pelaku bisnis dapat memperkirakan kemungkinan kejadian tertentu dan membuat keputusan berdasarkan probabilitas yang terukur. Misalnya, jika rata-rata pelanggan per jam adalah 4, maka distribusi Poisson bisa digunakan untuk memperkirakan kemungkinan lebih dari 5 pelanggan datang dalam satu jam.

f. Keterbatasan Distribusi Diskrit

- Tidak cocok digunakan jika variabel bersifat kontinu.
- Hanya bisa menangani kejadian-kejadian yang jumlahnya terbatas dan bisa dihitung.
- Hasil sangat bergantung pada asumsi model (seperti asumsi independensi dalam binomial).

g. Kesimpulan

Distribusi probabilitas diskrit menyediakan kerangka penting untuk memahami peristiwa-peristiwa yang dapat dihitung secara kuantitatif. Dalam konteks bisnis, distribusi ini memungkinkan manajer untuk merencanakan strategi berdasarkan estimasi probabilitas kejadian tertentu, seperti pembelian pelanggan, jumlah komplain, atau permintaan barang. Dengan pemahaman yang tepat, distribusi ini dapat membantu dalam mengoptimalkan operasi dan mengurangi risiko dalam pengambilan keputusan.

## 6.8 Distribusi Probabilitas Kontinu

Distribusi probabilitas kontinu digunakan untuk merepresentasikan variabel acak yang nilainya berada dalam rentang tak terhingga dan bersifat kontinu, artinya tidak dapat dihitung satu per satu. Contoh umum dari variabel ini termasuk tinggi badan seseorang, waktu yang dibutuhkan untuk menyelesaikan sebuah proses, atau besar penghasilan pelanggan dalam periode tertentu. Tidak seperti distribusi diskrit, pada distribusi kontinu, peluang untuk memperoleh nilai tunggal yang spesifik adalah nol. Oleh karena itu, analisis probabilitas dilakukan terhadap interval nilai, bukan pada titik nilai tertentu. Distribusi probabilitas kontinu ditentukan melalui fungsi kepadatan probabilitas (*Probability Density Function* atau PDF), yang secara matematis memberikan bentuk dari distribusi tersebut. Salah satu sifat mendasar dari PDF adalah bahwa luas area di bawah kurva fungsi ini untuk seluruh rentang nilai variabel acak harus sama dengan satu. Ini mencerminkan total probabilitas semua kemungkinan hasil.

a. Karakteristik Distribusi Probabilitas Kontinu
    1. Rentang Nilai yang Tak Terbatas dalam Suatu Interval
       Variabel acak yang bersifat kontinu mampu mengambil jumlah nilai yang tidak terbatas di antara dua titik pada suatu interval. Hal ini berarti bahwa di antara dua angka berapa pun, selalu ada kemungkinan nilai lainnya—misalnya, antara 1,5 dan 2,0 terdapat tak terhitung banyaknya nilai seperti 1,51; 1,501; 1,5001; dan seterusnya.
    2. Probabilitas titik tunggal adalah nol, sehingga probabilitas diukur pada suatu interval.
    3. Menggunakan fungsi densitas probabilitas (*Probability Density Function*/PDF) untuk

menggambarkan distribusi.

b. Fungsi Densitas Probabilitas (PDF)

Fungsi densitas probabilitas (f(x)) menggambarkan bagaimana nilai-nilai variabel acak didistribusikan.

Syarat-syarat:

- $f(x) \geq 0$ untuk semua x
- $\int f(x)\, dx = 1$ (Artinya, jika kita menjumlahkan seluruh kemungkinan probabilitas dari variabel acak kontinu di seluruh domainnya, totalnya harus 1 atau 100%.)

$$\int_{-\infty}^{\infty} f(x)\, dx = 1$$

Probabilitas bahwa variabel acak X berada dalam interval [a, b]: $P(a \leq X \leq b) = \int_a^b f(x)\, dx$

c. Distribusi Normal

Distribusi normal adalah distribusi paling penting dalam statistik, sering disebut sebagai "kurva lonceng" karena bentuknya yang simetris dan berbentuk lonceng. Salah satu distribusi kontinu yang paling penting dan sering digunakan dalam analisis ekonomi dan bisnis adalah distribusi normal. Distribusi ini memiliki bentuk lonceng simetris dan digunakan untuk menggambarkan banyak fenomena alam maupun sosial yang tersebar secara acak di sekitar rata-rata.

Karakteristik:

- Simetris terhadap rata-rata (mean)
- Mean = median = modus
- Ditetapkan oleh dua parameter: rata-rata ($\mu$) dan standar deviasi ($\sigma$)
- Luas total di bawah kurva adalah 1

Rumus fungsi kepadatan distribusi normal adalah sebagai berikut:

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

Keterangan:

- μ\muμ: nilai rata-rata (mean) dari distribusi,
- σ\sigmaσ: simpangan baku (standard deviation) yang mengukur penyebaran data dari rata-rata.

Distribusi normal digunakan dalam berbagai konteks seperti pengukuran kualitas produk, prediksi pendapatan pelanggan, hingga model risiko finansial. Distribusi normal standar memiliki μ = 0 dan σ = 1, dan digunakan dalam banyak aplikasi statistik dan bisnis.

d. Distribusi Uniform Kontinu

Distribusi uniform kontinu adalah tipe distribusi di mana setiap nilai dalam suatu rentang tertentu memiliki peluang yang sama untuk terjadi. Ini disebut juga sebagai distribusi seragam. Misalnya, jika waktu kedatangan pelanggan diukur dalam rentang antara pukul 10:00 hingga 11:00, dan semua waktu dalam interval tersebut dianggap sama kemungkinan, maka distribusinya bersifat uniform.

Rumus PDF untuk distribusi seragam kontinu antara aaa dan bbb adalah:

$$f(x) = \frac{1}{b-a} \quad \text{untuk } a \leq x \leq b$$

Distribusi ini sering dipakai ketika tidak ada preferensi terhadap suatu nilai dalam interval tertentu atau ketika asumsi awal bersifat netral.

e. Aplikasi Distribusi Kontinu dalam Bisnis dan Ekonomi

1. Analisis Risiko Keuangan: Estimasi return saham, harga obligasi, dan nilai tukar sering diasumsikan mengikuti distribusi normal.
2. Peramalan Permintaan: Prediksi jumlah penjualan atau permintaan barang dalam periode tertentu.
3. Kepuasan Pelanggan dan Waktu Tunggu: Analisis distribusi waktu tunggu pelanggan dalam antrean layanan.
4. Manajemen Produksi: Pengukuran variasi dalam ukuran atau kualitas produk manufaktur.

f. Distribusi Lain yang Relevan

- Distribusi Eksponensial: Untuk memodelkan waktu antara dua kejadian, misalnya waktu antar kedatangan pelanggan.
- Distribusi Uniform Kontinu: Semua nilai dalam suatu interval memiliki peluang yang sama. Contoh: waktu kedatangan antara 9:00–10:00.

g. Interpretasi dan Penggunaan Praktis

Distribusi kontinu memungkinkan analisis statistik yang lebih akurat terhadap fenomena yang melibatkan pengukuran kuantitatif berkesinambungan. Dalam bisnis, hal ini memungkinkan manajer membuat prediksi yang lebih baik mengenai tren dan fluktuasi data keuangan atau operasional.

Misalnya, jika diketahui bahwa rata-rata waktu pelayanan pelanggan adalah 10 menit dengan standar deviasi 2 menit, maka menggunakan distribusi normal, dapat dihitung berapa peluang waktu pelayanan melebihi 12 menit.

f. Kesimpulan

Distribusi probabilitas kontinu adalah alat penting dalam analisis statistik dan pengambilan keputusan bisnis. Dengan memahami karakteristik dan penerapan distribusi seperti normal dan eksponensial, pelaku bisnis dapat mengidentifikasi pola, mengukur risiko, dan menyusun strategi berdasarkan data yang bersifat kontinu. Ini memberikan keunggulan dalam perencanaan, pengawasan mutu, serta pengendalian proses dalam dunia bisnis modern.

## 6.9 Distribusi Sampling dan Aplikasinya dalam Ekonomi dan Bisnis

Distribusi sampling adalah distribusi probabilitas dari suatu statistik (seperti rata-rata atau proporsi) yang dihitung dari berbagai sampel acak berukuran tetap yang diambil dari populasi. Konsep ini penting karena memungkinkan pengambilan keputusan berbasis sampel untuk mewakili populasi.

a. Pengertian Distribusi Sampling

Distribusi sampling (*sampling distribution*) menjelaskan bagaimana nilai-nilai statistik seperti rata-rata sampel (mean), standar deviasi, atau proporsi akan terdistribusi jika dilakukan pengambilan sampel berkali-kali dari populasi yang sama.

Contoh: Jika kita mengambil 100 sampel acak dari populasi pelanggan dan menghitung rata-rata pengeluaran mereka, lalu ulangi proses tersebut 1.000 kali, maka kita dapat membentuk distribusi sampling dari rata-rata pengeluaran tersebut.

b. Statistik Sampel vs Parameter Populasi

- Parameter adalah nilai yang merangkum karakteristik seluruh populasi (contoh: rata-rata pendapatan semua pelanggan).
- Statistik adalah nilai yang dihitung dari sampel (contoh: rata-rata pendapatan 100 pelanggan yang disurvei).

Distribusi sampling menjelaskan bagaimana statistik mendekati parameter ketika ukuran sampel meningkat.

c. Teorema Limit Tengah (*Central Limit Theorem*/CLT)

Teorema Limit Tengah menyatakan bahwa:

- Jika ukuran sampel cukup besar ($n \geq 30$), maka distribusi sampling dari rata-rata sampel akan mendekati distribusi normal, terlepas dari bentuk distribusi populasi asalnya.
- Rata-rata distribusi sampling sama dengan rata-rata populasi: $\mu_{\bar{x}} = \mu$
- Standar deviasi dari distribusi sampling (disebut standar galat atau standard error) adalah: $\sigma_{\bar{x}} = \sigma / \sqrt{n}$

d. Aplikasi dalam Ekonomi dan Bisnis

1. Perkiraan Pendapatan Pelanggan: Menggunakan sampel pelanggan untuk memperkirakan rata-rata pendapatan semua pelanggan dengan estimasi galat tertentu.
2. Audit dan Pengendalian Kualitas: Mengambil sampel produk untuk memperkirakan proporsi cacat dalam produksi.
3. Survei Pasar: Menyimpulkan preferensi konsumen dari hasil survei sampel.
4. Pengambilan Keputusan Investasi: Mengestimasi return rata-rata suatu portofolio berdasarkan data

historis sampel.

e. Simulasi dan Visualisasi

Distribusi sampling dapat divisualisasikan melalui simulasi komputer, seperti dengan perangkat lunak Excel, R, atau Python. Misalnya, dengan membuat histogram dari ribuan rata-rata sampel untuk menunjukkan bentuk distribusi sampling.

f. Signifikansi Statistik dan Inferensi

Distribusi sampling merupakan dasar dari inferensi statistik. Dengan mengetahui distribusi sampling:

- Kita dapat melakukan uji hipotesis (*hypothesis testing*) untuk menyimpulkan apakah suatu pernyataan tentang populasi didukung oleh data.

Contoh:

- Menilai apakah rata-rata gaji karyawan sebuah perusahaan lebih tinggi dari standar industri.
- Menentukan apakah perubahan strategi pemasaran meningkatkan rata-rata pembelian pelanggan secara signifikan.

g. Kesimpulan

Distribusi sampling adalah pilar penting dalam statistik inferensial yang memungkinkan analis ekonomi dan pelaku bisnis membuat keputusan dengan dasar probabilistik dari data sampel. Dengan memahami cara kerja distribusi sampling dan aplikasinya, kita dapat membuat estimasi yang akurat, mengukur risiko, serta mengoptimalkan strategi bisnis dan kebijakan publik berdasarkan data yang representatif dan terpercaya.

## 6.10 Statistik dan Probabilitas dalam Pengambilan Keputusan Ekonomi

Statistik dan probabilitas adalah alat penting dalam pengambilan keputusan ekonomi karena mereka memungkinkan analis untuk mengukur ketidakpastian, memperkirakan hasil masa depan, dan membuat keputusan yang lebih terinformasi berdasarkan data yang tersedia. Keputusan dalam ekonomi, baik di sektor publik maupun swasta, sering kali melibatkan ketidakpastian dan risiko. Dalam konteks ini, statistik dan probabilitas menyediakan kerangka kerja untuk membuat prediksi yang lebih akurat dan mengelola risiko dengan lebih baik.

a. Peran Statistik dalam Keputusan Ekonomi

1. Analisis Data

   Statistik berfungsi sebagai alat untuk mengumpulkan, menganalisis, dan menginterpretasikan data. Dalam ekonomi, ini bisa berarti menganalisis data tentang pengeluaran konsumen, inflasi, atau produk domestik bruto (PDB). Dengan menggunakan teknik statistik yang tepat, data yang diperoleh dari survei atau studi pasar dapat dianalisis untuk memberikan wawasan yang relevan bagi pembuat kebijakan dan pengusaha.

2. Estimasi Parameter Populasi

   Sering kali, kita hanya dapat mengakses sampel data dan tidak seluruh populasi. Dalam hal ini, statistik memungkinkan kita untuk mengestimasi parameter populasi seperti rata-rata pendapatan atau tingkat pengangguran menggunakan data sampel yang diperoleh. Dengan

teknik estimasi, kita dapat menarik kesimpulan yang lebih luas tentang populasi dari sampel yang terbatas.

3. Analisis Tren dan Fluktuasi

   Statistik digunakan untuk menganalisis tren jangka panjang dan fluktuasi dalam perekonomian, misalnya dalam pengamatan terhadap pertumbuhan ekonomi, inflasi, dan pengangguran. Alat statistik seperti regresi linier atau analisis time-series membantu ekonom dan pembuat kebijakan untuk memahami dinamika pasar dan merumuskan kebijakan yang lebih tepat dan responsif terhadap perubahan ekonomi.

b. Peran Probabilitas dalam Keputusan Ekonomi

1. Pengukuran Risiko

   Ketika seorang investor memutuskan untuk membeli saham, mereka menggunakan probabilitas untuk memperkirakan kemungkinan kenaikan atau penurunan harga saham tersebut. Ini sangat penting dalam pengelolaan portofolio investasi dan perencanaan keuangan.

2. Peramalan Permintaan dan Penawaran

   Probabilitas juga digunakan dalam peramalan permintaan dan penawaran pasar. Sebagai contoh, dalam bisnis, probabilitas dapat digunakan untuk memperkirakan jumlah produk yang akan dibeli konsumen selama periode tertentu. Peramalan ini sangat penting untuk menentukan volume produksi dan pengelolaan persediaan barang.

3. Kebijakan Publik dan Pembangunan Ekonomi

   Pemerintah sering menggunakan probabilitas untuk menilai dampak kebijakan tertentu, seperti

insentif fiskal atau pajak, terhadap perekonomian. Dengan memodelkan berbagai skenario kemungkinan, probabilitas membantu pemerintah merencanakan kebijakan yang dapat meminimalkan risiko dan mencapai tujuan ekonomi yang diinginkan.

c. Model Pengambilan Keputusan

Biasanya, proses ini melibatkan beberapa langkah:

- Penilaian Keputusan

  Dengan menggunakan data historis dan model probabilistik, setiap alternatif keputusan dievaluasi berdasarkan kemungkinan dan hasil yang diharapkan. Misalnya, keputusan investasi dapat dievaluasi berdasarkan probabilitas pengembalian yang diharapkan.
- Analisis Sensitivitas

  Pengambil keputusan untuk memahami seberapa stabil hasil yang diperoleh dalam kondisi yang berbeda.
- Optimasi

  Pengambilan keputusan berbasis probabilitas juga dapat melibatkan optimasi, yang bertujuan untuk memilih alternatif yang memberikan hasil terbaik dengan memaksimalkan keuntungan atau meminimalkan kerugian. Dalam ekonomi, ini sering diterapkan dalam perencanaan produksi, pengelolaan sumber daya, atau perumusan kebijakan publik.

d. Aplikasi dalam Bisnis dan Ekonomi

1. Investasi dan Portofolio

   Dalam dunia investasi, probabilitas digunakan untuk menghitung risiko dan imbal hasil dari

berbagai aset dalam portofolio. Misalnya, pengelola portofolio dapat menggunakan teori portofolio modern dan probabilitas untuk memilih kombinasi aset yang memaksimalkan return yang diharapkan sambil meminimalkan risiko.

2. Peramalan Ekonomi

   Pemerintah dan perusahaan menggunakan statistik dan probabilitas untuk meramalkan kondisi ekonomi di masa depan. Peramalan ini sangat penting untuk merencanakan strategi bisnis, mengidentifikasi tren ekonomi, dan merumuskan kebijakan yang dapat merangsang pertumbuhan ekonomi.

3. Penentuan Harga

   Dalam bisnis, probabilitas digunakan untuk menentukan harga produk atau layanan. Misalnya, perusahaan dapat menganalisis data tentang harga produk serupa di pasar dan menghitung harga yang dapat memaksimalkan keuntungan sambil mempertimbangkan permintaan konsumen dan elastisitas harga.

4. Evaluasi Kebijakan Publik

   Pemerintah menggunakan statistik untuk mengevaluasi dampak kebijakan publik, seperti kebijakan perpajakan, subsidi, atau program bantuan sosial. Dengan menggunakan data dan probabilitas, mereka dapat memperkirakan efek dari kebijakan tertentu terhadap perekonomian, seperti dampaknya terhadap pengangguran atau pertumbuhan PDB.

e. Teknik-teknik Statistik yang Digunakan dalam Pengambilan Keputusan Ekonomi

1. Analisis Regresi

   Digunakan untuk mengidentifikasi hubungan antara variabel-variabel ekonomi, seperti antara pengeluaran pemerintah dan pertumbuhan ekonomi.

2. Uji Hipotesis

   Uji hipotesis digunakan untuk menguji teori-teori ekonomi atau mengonfirmasi asumsi yang digunakan dalam analisis ekonomi.

3. Analisis Varians (ANOVA)

   Teknik ini digunakan untuk membandingkan rata-rata dari dua atau lebih kelompok untuk melihat apakah ada perbedaan yang signifikan. Misalnya, ANOVA dapat digunakan untuk menganalisis apakah perbedaan dalam strategi pemasaran mempengaruhi penjualan produk secara signifikan.

f. Kesimpulan

Statistik dan probabilitas adalah fondasi penting dalam pengambilan keputusan ekonomi. Dengan memberikan alat untuk mengukur risiko, memperkirakan hasil, dan mengoptimalkan keputusan, kedua bidang ini membantu membuat keputusan yang lebih baik, baik di sektor swasta maupun publik. Penggunaan statistik dan probabilitas memungkinkan pembuat kebijakan, perusahaan, dan individu untuk membuat keputusan yang lebih terinformasi dan responsif terhadap dinamika ekonomi yang selalu berubah.

## 6.11 Pemahaman Probabilitas dalam Konteks Bisnis dan Ekonomi

Probabilitas, sebagai alat untuk mengukur ketidakpastian dan menganalisis risiko, sangat penting dalam pengambilan keputusan yang berkaitan dengan ekonomi dan bisnis. Dalam konteks ini, probabilitas digunakan untuk membantu berbagai aspek analisis dan perencanaan, dari manajemen risiko hingga evaluasi kinerja bisnis. Pemahaman probabilitas digunakan untuk: (i) Analisis risiko dan ketidakpastian, seperti pada asuransi dan investasi; (ii) Peramalan permintaan dalam manajemen persediaan; (iii) Analisis keputusan dalam strategi bisnis; (iv) Evaluasi kinerja, misalnya peluang keterlambatan pengiriman atau retur barang. Berikut adalah beberapa area utama di mana pemahaman probabilitas sangat berperan:

1. Analisis Risiko dan Ketidakpastian

   Salah satu aplikasi probabilitas yang paling mendalam adalah dalam analisis risiko dan ketidakpastian. Risiko dalam bisnis bisa datang dari berbagai sumber, baik yang dapat diprediksi maupun yang tidak dapat diprediksi. Pemahaman probabilitas memungkinkan pengusaha dan manajer untuk menilai kemungkinan terjadinya peristiwa tertentu serta dampaknya, sehingga mereka dapat membuat keputusan yang lebih baik dalam menghadapi ketidakpastian.

   a. Asuransi

      Di sektor asuransi, probabilitas digunakan untuk menghitung premi yang wajar dan untuk memprediksi kemungkinan kerugian yang harus dibayar kepada pemegang polis. Perusahaan asuransi menggunakan analisis probabilitas untuk

mengelola eksposur mereka terhadap risiko, dengan tujuan agar mereka dapat tetap menguntungkan meskipun ada klaim yang tinggi.

b. Investasi dan Portofolio

Dalam dunia investasi, probabilitas digunakan untuk menghitung potensi risiko dan imbal hasil. Melalui model seperti teori portofolio Modern, investor dapat menilai kemungkinan keuntungan atau kerugian yang akan diperoleh dari berbagai instrumen investasi, seperti saham, obligasi, atau komoditas. Pendekatan ini memungkinkan investor untuk memilih investasi yang optimal dengan mengurangi risiko dan memaksimalkan pengembalian yang diharapkan.

c. Evaluasi Kebijakan Bisnis

Dalam manajemen risiko perusahaan, probabilitas juga digunakan untuk menilai dampak dari berbagai keputusan strategis. Misalnya, perusahaan yang merencanakan ekspansi ke pasar baru dapat menggunakan probabilitas untuk mengukur kemungkinan sukses dan risiko yang terlibat. Hal ini membantu perusahaan untuk menentukan apakah langkah ekspansi tersebut sebanding dengan potensi risiko yang ada.

2. Peramalan Permintaan dalam Manajemen Persediaan

Manajemen persediaan adalah bagian penting dalam operasi bisnis, khususnya di sektor ritel dan manufaktur. Probabilitas digunakan untuk memprediksi permintaan produk di masa depan dan membantu perusahaan merencanakan persediaan yang cukup, tanpa terjebak dalam stok berlebihan atau kekurangan barang.

a. Peramalan Permintaan

Perusahaan menggunakan teknik probabilitas untuk meramalkan permintaan produk berdasarkan data historis dan variabel eksternal yang mempengaruhi pasar. Misalnya, perusahaan dapat menggunakan distribusi probabilitas untuk menghitung kemungkinan permintaan produk pada periode tertentu, berdasarkan tren musiman, peristiwa ekonomi, atau perilaku konsumen. Dengan informasi ini, manajer dapat memastikan bahwa mereka memiliki stok yang cukup untuk memenuhi permintaan tanpa menambah biaya penyimpanan yang berlebihan.

b. Pengelolaan Persediaan Berbasis Probabilitas

Dalam pengelolaan persediaan, probabilitas digunakan untuk menentukan tingkat persediaan yang optimal. Model probabilistik seperti *Economic Order Quantity* (EOQ) atau metode pemesanan ulang berbasis probabilitas membantu perusahaan meminimalkan biaya persediaan dengan memastikan bahwa mereka memiliki jumlah barang yang tepat pada waktu yang tepat. Ini memungkinkan perusahaan untuk mempertahankan keseimbangan antara ketersediaan produk dan biaya penyimpanan, serta mengurangi risiko kehabisan barang atau terlalu banyak barang yang terbuang.

3. Analisis Keputusan dalam Strategi Bisnis

Strategi bisnis yang efektif memerlukan kemampuan untuk mengevaluasi berbagai alternatif keputusan dalam menghadapi ketidakpastian. Probabilitas memberikan cara untuk menilai

kemungkinan hasil dari setiap keputusan dan dampaknya terhadap tujuan bisnis, baik dalam jangka pendek maupun jangka panjang.

a. Model Pengambilan Keputusan

Probabilitas digunakan dalam model pengambilan keputusan untuk menganalisis berbagai alternatif dan memilih strategi yang memberikan hasil terbaik dengan mempertimbangkan ketidakpastian. Misalnya, perusahaan yang merencanakan peluncuran produk baru dapat menggunakan analisis probabilitas untuk menilai kemungkinan keberhasilan peluncuran tersebut, baik dari sisi penerimaan pasar, biaya, dan persaingan.

b. Keputusan Investasi dan Strategi Ekspansi

Dalam pengambilan keputusan strategis, seperti ekspansi perusahaan ke pasar baru atau investasi dalam teknologi baru, probabilitas digunakan untuk memodelkan kemungkinan hasil yang berbeda. Ini termasuk analisis risiko yang terlibat dalam keputusan tersebut dan penilaian dampak dari variabel eksternal yang dapat mempengaruhi hasil, seperti perubahan ekonomi atau regulasi pemerintah. Dengan menggunakan probabilitas, perusahaan dapat memilih opsi yang memiliki kemungkinan terbesar untuk mencapai tujuan strategis mereka.

4. Evaluasi Kinerja dan Peluang Keterlambatan

Penerapan probabilitas dalam evaluasi kinerja bisnis membantu perusahaan untuk mengidentifikasi potensi masalah atau peluang yang mungkin tidak terlihat dalam analisis tradisional. Dalam konteks ini,

probabilitas digunakan untuk mengevaluasi berbagai indikator kinerja bisnis dan mengidentifikasi risiko atau peluang yang dapat memengaruhi efisiensi operasional.

a. Peluang Keterlambatan Pengiriman

Perusahaan yang mengelola rantai pasokan atau logistik menggunakan probabilitas untuk mengevaluasi kemungkinan keterlambatan pengiriman atau masalah dalam proses distribusi. Dengan menggunakan data historis dan model probabilistik, mereka dapat memprediksi kemungkinan terjadinya keterlambatan dalam pengiriman barang, misalnya karena cuaca buruk atau masalah pada sistem transportasi. Ini memungkinkan perusahaan untuk mengambil langkah-langkah proaktif, seperti mengatur cadangan pengiriman atau memperbaiki sistem logistik, untuk mengurangi dampak negatif dari keterlambatan.

b. Retur Barang dan Ketidakpuasan Pelanggan

Probabilitas juga digunakan untuk memodelkan kemungkinan pengembalian barang atau ketidakpuasan pelanggan yang dapat memengaruhi kinerja bisnis. Dengan menggunakan analisis probabilitas, perusahaan dapat memperkirakan tingkat pengembalian barang dan menentukan langkah-langkah untuk mengurangi jumlah retur, seperti memperbaiki kualitas produk atau memberikan penjelasan yang lebih baik kepada pelanggan tentang penggunaan produk. Ini membantu perusahaan untuk meningkatkan pengalaman pelanggan dan mengurangi biaya terkait retur barang.

b. Pengukuran Kinerja Operasional

Perusahaan menggunakan probabilitas untuk mengevaluasi kinerja operasional mereka dalam berbagai aspek, seperti kecepatan produksi, pengelolaan persediaan, dan biaya operasional. Dengan menghitung probabilitas terjadinya kegagalan dalam sistem operasional, manajer dapat mengidentifikasi area yang membutuhkan perbaikan dan mengambil langkah-langkah untuk meningkatkan efisiensi dan produktivitas.

5. Kesimpulan

Pemahaman probabilitas dalam ekonomi dan bisnis memungkinkan perusahaan untuk menghadapi ketidakpastian dan risiko dengan lebih baik. Dengan mengaplikasikan probabilitas dalam analisis risiko, peramalan, pengambilan keputusan strategis, dan evaluasi kinerja, perusahaan dapat membuat keputusan yang lebih tepat dan meningkatkan efisiensi operasional mereka. Secara keseluruhan, probabilitas memberikan dasar yang kuat untuk perencanaan dan pengelolaan risiko yang lebih baik, serta memungkinkan perusahaan untuk beradaptasi dengan perubahan yang cepat di pasar dan lingkungan ekonomi.

## 6.12 Penerapan Probabilitas dalam Ekonomi dan Bisnis

Probabilitas memainkan peran penting dalam pengambilan keputusan yang berkaitan dengan ekonomi dan bisnis, mengingat ketidakpastian yang melekat dalam banyak aspek kehidupan ekonomi. Dalam konteks ekonomi dan bisnis, probabilitas digunakan untuk mengukur ketidakpastian, memprediksi kemungkinan hasil dari

keputusan, serta mengelola risiko. Aplikasi probabilitas dalam bidang ini mencakup berbagai situasi, dari peramalan pasar hingga evaluasi investasi dan kebijakan ekonomi. Dengan demikian, pemahaman yang baik tentang probabilitas menjadi kunci untuk membuat keputusan yang lebih informatif dan efektif.

a. Peramalan Ekonomi dan Bisnis

Salah satu penerapan utama probabilitas dalam ekonomi dan bisnis adalah dalam peramalan, di mana probabilitas digunakan untuk memprediksi kejadian-kejadian di masa depan; meliputi beberapa variabel ekonomi, seperti inflasi, pengangguran, harga barang, dan pertumbuhan ekonomi, dan lain-lain.

1. Peramalan Pertumbuhan Ekonomi

Perusahaan dan pemerintah menggunakan probabilitas untuk memprediksi kondisi ekonomi yang akan datang, seperti laju pertumbuhan ekonomi atau fluktuasi pasar. Dengan menganalisis data historis dan menggunakan model probabilistik, mereka dapat mengidentifikasi tren ekonomi dan mempersiapkan langkah-langkah strategis yang lebih baik.

2. Peramalan Permintaan Pasar

Dalam bisnis, probabilitas digunakan untuk memprediksi permintaan produk atau layanan di masa depan. Misalnya, perusahaan ritel dapat menggunakan probabilitas untuk memperkirakan berapa banyak produk yang akan terjual dalam periode tertentu berdasarkan data historis, tren pasar, dan faktor musiman. Ini memungkinkan perusahaan untuk merencanakan produksi dan persediaan dengan lebih efisien.

b. Pengelolaan Risiko dalam Bisnis

Risiko adalah bagian integral dari setiap keputusan bisnis. Probabilitas membantu dalam pengelolaan risiko dengan memberikan alat untuk menilai kemungkinan dan dampak dari berbagai risiko yang dapat mempengaruhi operasi bisnis.

1. Evaluasi Investasi dan Portofolio

   Dalam bisnis terdapat beberapa penerapan utama probabilitas dimana evaluasi investasi adalah salah satunya. Investor menggunakan probabilitas untuk menghitung risiko dan imbal hasil dari berbagai jenis investasi. Dalam model portofolio, probabilitas digunakan untuk memaksimalkan pengembalian yang diharapkan sambil meminimalkan risiko. Ini termasuk analisis probabilistik dari variasi harga saham, obligasi, atau komoditas.

2. Asuransi

   Dalam industri asuransi, probabilitas digunakan untuk menghitung premi yang harus dibayar oleh pemegang polis untuk berbagai jenis perlindungan. Perusahaan asuransi menggunakan probabilitas untuk memperkirakan kemungkinan terjadinya klaim dan dampaknya terhadap keuangan perusahaan. Oleh karena itu, probabilitas memungkinkan perusahaan asuransi untuk menetapkan harga polis yang adil dan mengelola risiko yang terkait dengan klaim.

c. Pengambilan Keputusan dalam Investasi

Keputusan investasi sering kali melibatkan ketidakpastian, dan probabilitas memberikan alat untuk menilai risiko serta potensi keuntungan dari berbagai

pilihan investasi.

1. Keputusan Investasi Saham

   Dalam dunia investasi saham, probabilitas digunakan untuk menilai kemungkinan harga saham tertentu akan naik atau turun dalam jangka waktu tertentu. Dengan menggunakan model probabilistik, investor dapat memperkirakan tingkat pengembalian yang diharapkan dan risiko terkait, yang kemudian dapat digunakan untuk mengoptimalkan keputusan pembelian atau penjualan saham.

2. Evaluasi Proyek Bisnis

   Dalam bisnis, evaluasi proyek investasi atau bisnis baru sering melibatkan analisis probabilistik. Misalnya, dalam memutuskan apakah akan meluncurkan produk baru, perusahaan akan menggunakan probabilitas untuk memperkirakan peluang keberhasilan proyek tersebut, termasuk proyeksi pendapatan, biaya, dan laba yang diharapkan. Ini memungkinkan perusahaan untuk memilih proyek yang memiliki probabilitas tertinggi untuk memberikan keuntungan.

d. Penerapan dalam Analisis Kebijakan Ekonomi

Probabilitas juga digunakan untuk merumuskan dan mengevaluasi kebijakan ekonomi, baik itu kebijakan fiskal, moneter, atau kebijakan perdagangan internasional.

1. Evaluasi Kebijakan Fiskal dan Moneter

   Pemerintah menggunakan probabilitas untuk mengevaluasi dampak kebijakan fiskal dan moneter. Sebagai contoh, ketika pemerintah

meningkatkan pengeluaran atau mengubah tarif pajak, probabilitas dapat digunakan untuk memodelkan dampak perubahan ini terhadap pertumbuhan ekonomi, inflasi, atau pengangguran. Dengan menggunakan model probabilistik, pembuat kebijakan dapat memperkirakan kemungkinan dampak positif atau negatif dari kebijakan yang diusulkan.

2. Kebijakan Perdagangan Internasional

   Negara juga menggunakan probabilitas untuk menilai potensi risiko dan manfaat dari kebijakan perdagangan internasional. Misalnya, ketika negara-negara mempertimbangkan untuk mengadopsi tarif perdagangan tertentu atau membuka pasar untuk produk tertentu, probabilitas digunakan untuk memprediksi dampak ekonomi dari kebijakan tersebut terhadap perdagangan, investasi asing, dan lapangan kerja.

e. Teknik-Teknik Probabilitas dalam Bisnis dan Ekonomi

1. Distribusi Normal

   Distribusi normal sering digunakan dalam analisis ekonomi dan bisnis untuk menggambarkan distribusi variabel acak, seperti harga saham atau tingkat pengangguran. Dengan memahami distribusi normal, analis dapat mengukur kemungkinan pergerakan harga atau variabel lain dalam kisaran tertentu dan mengidentifikasi peluang atau ancaman.

2. Analisis Regresi

   Teknik regresi digunakan untuk memodelkan hubungan antara satu atau lebih variabel independen dan variabel dependen. Dalam

ekonomi, analisis regresi sering digunakan untuk memprediksi hubungan antara faktor-faktor seperti suku bunga, inflasi, dan tingkat pengangguran. Dengan menggunakan model regresi, analisis probabilitas dapat digunakan untuk memperkirakan hasil ekonomi di masa depan.

3. Simulasi Monte Carlo

   Teknik Simulasi Monte Carlo dipakai mengetahui dampak ketidakpastian dalam model ekonomi dan bisnis yang melibatkan penggunaan distribusi probabilitas untuk menghasilkan hasil yang berbeda dari suatu model, memungkinkan pembuat keputusan untuk mengevaluasi berbagai hasil yang mungkin terjadi dalam berbagai kondisi.

f. Kesimpulan

   Penerapan probabilitas dalam ekonomi dan bisnis memberikan dasar yang kuat untuk pengambilan keputusan yang lebih baik dan lebih informatif., Organisasi/ lembaga/ Perusahaan, pemerintah, dan individu dapat membuat keputusan yang lebih terukur dan berkelanjutan setelah mengetahui keberadaan ketidakpastian dan risiko. Dari peramalan ekonomi dan investasi hingga kebijakan publik dan manajemen risiko, probabilitas memungkinkan pengambilan keputusan yang lebih efisien dan efektif dalam menghadapi tantangan ekonomi yang kompleks.

## 6.13 Kesimpulan dan Penutup

Probabilitas merupakan fondasi penting dalam analisis statistik, terutama untuk memahami dan mengelola ketidakpastian dalam dunia ekonomi dan bisnis. Dengan memanfaatkan konsep-konsep seperti kejadian, ruang

sampel, probabilitas bersyarat, dan aturan probabilitas, pengambil keputusan dapat melakukan penilaian risiko yang lebih rasional dan terukur.

Dalam bab ini, kita telah membahas berbagai konsep dasar dalam statistik dan probabilitas yang sangat penting dalam konteks ekonomi dan bisnis. Probabilitas, sebagai alat untuk mengukur ketidakpastian dan risiko, memainkan peran yang sangat vital dalam pengambilan keputusan yang berbasis data, baik dalam perencanaan, manajemen risiko, strategi bisnis, maupun evaluasi kinerja.

Probabilitas memberikan kerangka kerja yang kokoh untuk memahami berbagai kemungkinan hasil yang dapat terjadi di masa depan, yang memungkinkan perusahaan untuk merencanakan lebih baik dan mengambil langkah-langkah mitigasi yang tepat. Konsep dasar seperti distribusi probabilitas, probabilitas bersyarat, serta berbagai teorema dan pendekatan statistik yang telah dibahas, sangat relevan untuk mendukung analisis dan perencanaan dalam dunia bisnis.

Beberapa penerapan praktis probabilitas yang dibahas, antara lain dalam analisis risiko, peramalan permintaan, pengambilan keputusan dalam strategi bisnis, dan evaluasi kinerja, menunjukkan bahwa pemahaman probabilitas tidak hanya terbatas pada teori, tetapi juga memiliki dampak langsung terhadap efisiensi dan efektivitas operasi bisnis. Oleh karena itu, kemampuan untuk menggunakan probabilitas secara tepat dan bijaksana menjadi suatu kompetensi yang sangat dibutuhkan oleh para pengambil keputusan dan profesional di bidang ekonomi dan bisnis.

Secara keseluruhan, statistik dan probabilitas menyediakan alat yang esensial bagi para profesional untuk mengelola ketidakpastian dan mengambil keputusan yang lebih terinformasi. Dengan pemahaman yang kuat tentang

konsep-konsep probabilitas ini, baik perusahaan kecil maupun besar dapat mengembangkan strategi yang lebih baik, memitigasi risiko, dan meningkatkan daya saing di pasar yang semakin kompetitif.

Sebagai penutup, penerapan yang tepat dari statistik dan probabilitas dalam ekonomi dan bisnis tidak hanya meningkatkan kinerja perusahaan, tetapi juga memberikan landasan yang lebih baik untuk mencapai tujuan jangka panjang yang berkelanjutan. Oleh karena itu, pembelajaran dan penerapan konsep-konsep ini akan terus relevan dan berkembang seiring dengan dinamika pasar dan kemajuan teknologi yang terjadi.

Dengan demikian, buku ini diharapkan dapat memberikan wawasan yang mendalam tentang pentingnya statistik dan probabilitas dalam pengambilan keputusan ekonomi dan bisnis, serta membantu pembaca untuk lebih memahami bagaimana konsep-konsep tersebut dapat diterapkan dalam dunia nyata.

## DAFTAR PUSTAKA

Musnaini, Ida Ketut Mudhita, Asrini, Ida Ayu Widia Laksmi. (2023). Statistik Ekonomi: Metode Probabilitas. Sidoarjo: Indomedia Pustaka. ISBN 978-623-414-150-4.

Kuncoro, M. (2001). Metode Kuantitatif: Teori dan Aplikasi untuk Bisnis dan Ekonomi. Yogyakarta: AMP YKPN.

Anggita, K. T. (2022). Pengaruh Profitabilitas, Likuiditas, dan Leverage Terhadap Nilai Perusahaan. Journal of Economics and Business UBS, 12(4), 2087–2099.

Dewi, N. S., & Suwarno, A. E. (2022). Pengaruh ROA, ROE, EPS, dan DER Terhadap Harga Saham Perusahaan. Seminar Nasional Pariwisata Dan Kewirausahaan (SNPK), 1, 472–482.

Girsang, A. N., et al. (2019). Analisis Pengaruh EPS, DPR, dan DER terhadap Harga Saham Sektor Trade, Services, & Investment di BEI. Jesya (Jurnal Ekonomi & Ekonomi Syariah), 2(2), 351–362.

Badan Pusat Statistik (BPS). (2023). Statistik Indonesia 2023.

BPS Kabupaten Asahan. (2021). Statistik Gender Sumatera Utara Tahun 2016.

BPS Kabupaten Pesisir Selatan. (2022). Pesisir Selatan Dalam Angka.

Mediana. (2023, September 27). Permendag No 31/2023 Tetapkan 6 Model Bisnis Perdagangan Daring. Kompas.id.

Makki, S. (2023, October 11). Cerita Herni Omzet Melesat Rp2,5 Juta Usai TikTok Shop Dilarang Jualan. CNN Indonesia.

Maryoto, A. (2023, September 28). Pelajaran dari Kehadiran

"Social Commerce". Kompas.id.

Ghozali, Imam. (2018). Aplikasi Analisis Multivariate dengan Program IBM SPSS 25 Edisi 9. Semarang: Universitas Diponegoro.

Gujarati, D. N. (2015). Dasar-Dasar Ekonometrika Buku I Edisi Kelima.

Boediono. (1985). Ekonomi Moneter, Edisi 3. Yogyakarta: BPFE.

Sujarweni, V.W. (2019). Statistik untuk Bisnis dan Ekonomi.

# BAB 7
# DISTRIBUSI PROBABILITAS

**Oleh Joni Wilson Sitopu**

## 7.1 Pendahuluan

Bab ini membahas konsep ekspektasi, varians, serta fungsi kepadatan probabilitas pada variabel acak, baik yang bersifat diskrit maupun kontinu. Variabel acak merupakan variabel yang nilai numeriknya ditentukan berdasarkan peluang atau probabilitas. Pada umumnya, variabel acak terbagi menjadi dua tipe utama yang mudah diidentifikasi, yakni variabel diskrit dan variabel kontinu. Variabel diskrit hanya dapat mengambil sejumlah nilai tertentu yang terbatas, sedangkan variabel kontinu dapat memiliki nilai dalam rentang yang tidak terbatas. Fungsi kepadatan probabilitas dapat digunakan untuk menggambarkan perilaku dari beberapa jenis variabel acak diskrit dan kontinu. Dalam bab ini, dibahas sejumlah distribusi probabilitas penting untuk variabel acak diskrit, seperti distribusi Binomial, Hipergeometrik, Multinomial, Geometrik, Poisson, dan Binomial Negatif. Sementara itu, untuk variabel acak kontinu, dibahas berbagai distribusi probabilitas seperti distribusi Normal, Uniform, Eksponensial, Erlang, Gamma, Beta, Weibull, Lognormal, Student-t, F, dan Chi-Square ($\chi^2$).(dalam Zul Fadli Rr, Suprantiningrum, Joni Wilson Sitopu.,dkk.,(2023))

Pada Distribusi Binomial, Misalkan kita melakukan percobaan seperti melempar koin atau dadu berulang kali atau memilih kelereng dari guci berulang kali. Setiap lemparan atau pemilihan disebut percobaan. Dalam setiap percobaan tunggal akan ada probabilitas yang terkait dengan kejadian tertentu seperti gambar kepala pada koin, dadu, atau pemilihan kelereng merah. Dalam beberapa kasus, probabilitas ini tidak akan berubah dari satu percobaan ke percobaan berikutnya (seperti dalam melempar koin atau dadu). Percobaan tersebut kemudian dikatakan independen dan sering disebut percobaan Bernoulli setelah James Bernoulli yang menyelidikinya pada akhir abad ketujuh belas. (dalam Joni Wilson Sitopu, (2022)).

## 7.2 Distribusi Probabilitas

Konsep probabilitas berhubungan dengan percobaan yang hasilnya tidak dapat dipastikan sebelumnya. Dengan kata lain, meskipun percobaan dilakukan berulang kali dalam kondisi yang serupa, hasil yang muncul bisa saja berbeda setiap kali. Distribusi probabilitas berperan sebagai model yang digunakan untuk mempelajari hasil dari percobaan acak yang nyata, serta untuk memperkirakan kemungkinan hasil-hasil yang mungkin terjadi. Pernyataan ini berlaku apabila memenuhi syarat: (1) peristiwa yang terjadi dalam kenyataan merepresentasikan peristiwa yang dijelaskan oleh model, dan (2) kondisi-kondisi dalam model distribusi dapat direalisasikan dalam pelaksanaan percobaan tersebut.(dalam Sitopu. Joni Wilson, dkk,. (2022), Joni Wilson Sitopu,. (2023)).

**Definisi :** Distribusi probabilitas mencakup himpunan nilai yang mungkin diambil oleh suatu variabel acak, beserta probabilitas yang terkait dengan masing-masing nilai tersebut.

**Contoh:**

Perhatikan percobaan pelempar koin tiga kali.

Misalkan X adalah jumlah sisi kepala.

Buatlah distribusi probabilitas X.

Solusi:

1. Pertama-tama identifikasikan nilai yang mungkin dimiliki oleh X.
2. Hitunglah probabilitas dari setiap nilai X yang mungkin dan nyatakan X dalam bentuk distribusi frekuensi.

| X=x | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| P(X=x) | 1/8 | 3/8 | 3/8 | 1/8 |

Probabilitas distribusi dilambangkan dengan P untuk variabel acak diskrit dan dengan f untuk variabel acak kontinyu. Kapur, J.N dan Saxena, H.C. (1976)

**Sifat-sifat Distribusi Probabilitas :**

1. P(x)≥0, Jika X adalah diskrit
   f(x)≥0, jika X adalah kontinu

2. $\sum_x P(X = X) = 1,\ Jika\ X\ adalah\ diskrit.$
   $\int_x f(x)dx = 1$, jika x adalah kontinu

Dimana ;

1. jika X adalah variable random kontinu, maka
   $P(a < X < b) = \int_a^b f(x)dx$

2. Probabilitas suatu nilai tetap dari variabel random kontinyu adalah nol.

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b)$$

3. Jika X adalah variabel acak diskrit,
   $P(a < X < b) = \sum_{x=a+1}^{b-1} P(X)$
   $P(a \leq X < b) = \sum_{a}^{b-1} P(X)$
   $P(a < X \leq b) = \sum_{a+1}^{b} P(X)$
   $P(a \leq X \leq b) = \sum_{a}^{b} P(X)$
4. Probabilitas berarti luas untuk variabel acak kontinu.

### 7.2.1 Ekspektasi

**Definisi :**

1. Misalkan variabel random diskrit X memiliki nilai $X_1, X_2, \dots, X_n$ dengan probabilitas $P(X_1) + P(X_2) + \cdots + P(X_n)$ berturut-turut. Maka nilai harapan X, dilambangkan sebagai E(X) didefinisikan sebagai
   $E(X) = X_1P(X_1) + X_2P(X_2) + \cdots + X_nP(X_n)$
   $E(X) = \sum_{i=1}^{n} X_iP(X_i)$. Kapur, J.N dan Saxena, H.C. (1976)

2. Misalkan X adalah variabel randrom kontinu yang mengasumsikan nilai-nilai dalam interval (a, b) sehingga,
   $\int_a^b f(X)dx = 1$, maka $E(X) = \int_a^b Xf(X)dx$

### 7.2.2 Rata-rata dan Varians Dari Variabel Random

Misalkan X adalah random variable.

1. Nilai harapan X adalah nilai rata-ratanya
   $\bar{X} = E(X) = \mu$

2. Varians X adalah,
   Var(X) = $Var(X) = E(X^2) - (E(X))^2$
   Dimana ;

$$E(X^2) = \sum_{j=1}^{n} X_j^2 P(X = x_j)$$ , jika X diskrit

$$E(X^2) = \int_x X^2 f(x)dx$$, jika X kontinu

**Contoh :**

Tentukan mean dan varians sebuah variabel acak dengan distribusi Bernoulli berparameter p, yaitu X mempunyai p.d.f. sebagai berikut.

| X | 0 | 1 |
|---|---|---|
| f(x) | 1-p | p |

Solusi,

E(x) = 0.(1-p)+1.p = p

Var(x) = p-$p^2$ = p(1-p)

## 7.3 Distribusi Probabilitas Diskrit

Distribusi Binomial, Misalkan kita melakukan percobaan seperti melempar koin atau dadu berulang kali atau memilih kelereng dari guci berulang kali. Setiap lemparan atau pemilihan disebut percobaan. Dalam setiap percobaan tunggal akan ada probabilitas yang terkait dengan kejadian tertentu seperti gambar kepala pada koin, dadu, atau pemilihan kelereng merah. Dalam beberapa kasus, probabilitas ini tidak akan berubah dari satu percobaan ke percobaan berikutnya (seperti dalam melempar koin atau dadu). Percobaan tersebut kemudian dikatakan independen dan sering disebut percobaan Bernoulli setelah James Bernoulli yang menyelidikinya pada akhir abad ketujuh belas.

Distribusi binomial merupakan distribusi probabilitas diskrit yang menggambarkan jumlah terjadinya suatu peristiwa dalam sejumlah n percobaan acak yang bersifat

independen, di mana setiap percobaan memiliki peluang keberhasilan sebesar p. Dengan demikian, interpretasinya dapat berupa sebagai berikut:

- ✓ Variabel Acak: Jumlah kemunculan (X) dalam serangkaian n percobaan independen.
- ✓ Parameter: n dan p, di mana n adalah jumlah percobaan dan p adalah probabilitas kemunculan dalam setiap percobaan.

Fungsi Kepadatan Probabilitas ;

f(x) = $P(X=x) = \binom{n}{x} p^x(1-p)^{n-x}$, x = 0,1,...,n

Kapur, J.N dan Saxena, H.C. (1976)

mean = (E(X)=μ), varians ($\sigma^2$) distribusi binomial adalah sebagai berikut ;

mean = E(X)=μ = np

varians = $\sigma^2$= npq

**Contoh :**

Carilah rata-rata kejadian banjir 10 tahun (periode ulang banjir adalah 10 tahun) dalam periode 100 tahun? Berapa probabilitas bahwa tepat jumlah banjir 10 tahun ini akan terjadi dalam periode 100 tahun?

**Solusi :**

Peluang terjadinya banjir 10 tahun sekali pada tahun berapa pun = 1/10 = 0,1.

Jadi, rata-rata jumlah kejadian dalam 100 tahun = E(X) =μ= np = 100 × 0,1 = 10.

Peluang terjadinya banjir 10 tahun sebanyak 10 kali dalam 100 tahun dapat dievaluasi menggunakan distribusi binomial.

$P(X=x) = \binom{n}{x} p^x(1-p)^{n-x} = \binom{10}{100}(0,1)^{10}(0,9)^{90} = 0.1319$

Tabel berikut merupakan rangkuman mean, varians dan fungsi kepadatan probabilitas variable acak diskrit.

**Tabel 7. 1 Distribusi Probabilitas Diskrit**

| No | Distribusi | Model | | |
|---|---|---|---|---|
| | | Probabilitas | Mean | Varians |
| 1. | Binomial | f(x) = $P(X=x) = \binom{n}{x} p^x(1-p)^{n-x}$, x = 0,1,...,n | np | npq |
| 2 | Hipergeometrik | P(X=x) = $pq^x$, x = 0,1,2,... | np | $\frac{npq(N-n)}{N-1}$ |
| 3 | Multinomial | $P(X_1 = n_1, X_2 = n_2, \dots, X_k = n_k) = \frac{n}{n1!n2!\dots nk!} p_1^n p_k^{n1} p_k^{n2} \dots p_k^{nk}$ | | |
| 4 | Geometric | P(X=x) = $pq^x$, x = 0,1,2,..., | $\frac{q}{p}$ | $\frac{q}{p^2}$ |
| 5 | Binomial Negatif | $P(X = x) = \binom{x+n-1}{n-1}$ $p^n$ $q^x$; x = 0,1,2,... | $\frac{nq}{p}$ | $\frac{nq}{p^2}$ |
| 6 | Poisson | P(X=x) = $\frac{e^x e^{-\lambda}}{x!}$ ; x = 0,1,2,... | λ | λ |

*Sumber :* (dalam Joni Wilson Sitopu dan Siswadi, (2022)).

## 7.4 Distribusi Probabilitas Kontinu

Salah satu contoh paling penting dari distribusi probabilitas kontinu adalah distribusi normal, yang terkadang disebut distribusi Gaussian. Fungsi kepadatan untuk distribusi ini diberikan oleh,

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

di mana yang satu merepresentasikan rata-rata, dan lainnya adalah simpangan baku. Fungsi distribusi yang sesuai diberikan oleh,

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{1}{2}\right)\left(\frac{v-\mu}{\sigma}\right)^2} dv$$

Jika variabel acak X memiliki fungsi distribusi F(x), maka X disebut berdistribusi normal dengan parameter mean μ dan varians $\sigma^2$.

Jika Z adalah variabel standar dengan X, yaitu, misalkan, $Z = \frac{X-\mu}{\sigma}$

dengan demikian, nilai rata-rata atau nilai harapan dari Z adalah 0, dan variansinya adalah 1. Dalam kondisi tersebut, fungsi kerapatan probabilitas untuk Z dapat diturunkan dari persamaan f(x) yang telah disebutkan sebelumnya, dengan menggantikan nilai rata-rata dan variansi secara formal menjadi 0 dan 1, sehingga diperoleh:

$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ disebut fungsi kepadatan normal standar. (dalam Joni Wilson Sitopu., Nanang.,dkk.,(2023))

**Contoh :**

Jika X mempunyai distribusi normal standar yaitu mean X, μ=0, dan varians = 1. Carilah P(X ≤ 0.5)

**Solusi :**

$$P(X = x) = f(x) = \int_{-\sim}^{\sim} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx = F(0{,}5)$$

Dari tabel F(0.5) ditemukan dengan mencari dua digit pertama (0.5) pada kolom yang diberi judul x, probabilitas

yang diinginkan sebesar 0.6915 ditemukan pada baris yang diberi label 0.5 dan yang diberi label 0.00. Yaitu ;

$$P(X = x) = f(x) = \int_{-\sim}^{0,5} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}\, dx = 0.6915$$

**Contoh :**

Jika x adalah variabel normal μ = 5 dan varians $\sigma^2$ = 16, carilah; P(X) ≤ 6)

Solusi

$$P(X) \leq 6 = P(\frac{x-\mu}{\sigma} \leq \frac{6-5}{4} \leq \frac{1}{4} = P\left(Y \leq \frac{1}{4}\right) = 0.5987$$

**Contoh :**

Misalkan X adalah variabel acak binomial dengan parameter n = 40, dan peluang keberhasilan p = 0,20. Hitunglah probabilitas terjadinya peristiwa ketika X = 5.

**Solusi:**

Karena X bersifat diskrit dan distribusi normal adalah untuk variabel acak kontinu, kita tuliskan

$$P(X = 5) = P(4{,}5 < X < 5{,}5) = P(\frac{4{,}5-4}{1{,}90} < \frac{X-4}{1{,}90} > \frac{5{,}5-4}{1{,}90}$$

$$p(0{,}26 < y < 0{,}79) = \Phi(0{,}79 - \Phi(0{,}26) = 0.18226.$$

Tabel berikut merupakan rangkuman mean, varians dan fungsi kepadatan probabilitas variable acak kontinu.

**Tabel 7. 2 Distribusi Probabilitas Kontinu**

| No | Distribusi | Model | | |
|---|---|---|---|---|
| | | Probabilitas | Mean | Varians |
| 1. | Normal | $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{1}{2}\right)\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty \leq x \leq \infty$ л=3,14159… , | μ<br>μ=0 | $\sigma^2$<br>$\sigma^2$=1 |

| No | Distribusi | Model | | |
|---|---|---|---|---|
| | | Probabilitas | Mean | Varians |
| | | e=2,71828…<br>$Z = \frac{x-\mu}{\sigma}$ | | |
| 2 | Uniform | $f(x) = \frac{1}{b-a}$,<br>$jika\ a \leq x \leq b, dan\ 0\ lainnya$ | $\mu = \frac{a+b}{2}$ | $\sigma^2 = \frac{(b-a)^2}{12}$ |
| 3 | Eksponensial | $f(x) = \frac{1}{\lambda} e^{-\frac{\lambda}{x'}}\ \ x > 0, \lambda > 0$, dimana $\lambda > 0$ dan $\lambda$ adalah distribusi parameter | $\lambda$ | $\lambda^2$ |
| 4 | Erlang | $f(x) = \lambda e^{\frac{\lambda}{x}} \frac{(\lambda x)^{n-1}}{(n-1)!}$, $0 < x < \infty, dan\ 0\ untuk\ lainnya$, $\lambda > 0$, dan $\lambda$ adalah distribusi parameter. | | |
| 5 | Gamma | $f(x) = \frac{1}{\tau(\alpha)\beta^\alpha} X^{\alpha-1} e^{-\frac{x}{\beta}}$,<br>$X > 0, \alpha > 0, \beta > 0$ | αβ | αβ² |
| 6 | Beta | $f(X) = \frac{1}{B(\alpha,\beta)} X^{\alpha-1}(1-X)^{\beta-1}$,<br>$0 < X < 1, \alpha > 0, \beta > 0$ | $\frac{\alpha}{\alpha+\beta}$ | $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta}$ |
| 7 | Weibull | $f(y) = \frac{\alpha}{\theta} y^{\alpha-1} e^{-\frac{y^\alpha}{\theta}}, y > 0, \alpha > 0, \theta > 0$ | $\theta^{\frac{1}{\alpha}} \tau(1 + \frac{1}{\alpha})$ | $\theta^{\frac{2}{\alpha}} \tau(\tau\left(1 + \frac{2}{\alpha}\right) - ((1 + \frac{1}{\alpha})^2))$ |
| 8 | Lognormal | $f(x) = \frac{1}{\sqrt{2\pi}\sigma} X^{-1} e^{\frac{-(lnX-\mu)^2}{2\sigma^2}}$, | $e^{\mu+\frac{1}{2}\sigma^2}$ | $(e^{\sigma^2} - 1)e^{2\pi+e^2}$ |

| No | Distribusi | Model | | |
|---|---|---|---|---|
| | | Probabilitas | Mean | Varians |
| | | $-\infty < x < \infty,$ $-\infty < \mu < \infty,\ \sigma^2 > 0$ | | |
| 9 | Student (t) | $X = \frac{Z}{\sqrt{U/V}} \sim t_v$, maka $n \to \infty, X \to N(0,1)$ | 0 | 1 |
| 10 | F | $X = \frac{U/v_1}{V/v_2} \sim F_{v_1,v_2}$ | $\frac{v_2}{v_2 - 2'}$ $(v_2 > 2)$ | $\frac{2v_2^2(v_1 + v_2 -}{v_1(v_2 - 4)(v_2 -}$ $(v_2 > 4)$ |
| 11 | Chi Square ($X^2$) | $f(x) = \frac{1}{\tau(\frac{\alpha}{2})2^{\frac{\alpha}{2}}} X^{\frac{\alpha}{2}-1} e^{-\frac{x}{2}},$ $X > 0, \alpha > 0$ | α | 2α |

Sumber : Joni Wilson Sitopu, (1990).

## DAFTAR PUSTAKA

Hogg, Robert V, Allen T. (1978)., *Introduction to Mathematical Statistics*, Fourth Edition, Macmillan Publishing co., inc., New York.

I Nyoman Susila dan Elle Gunawan, 1994. *Statistika*. Penerbit Erlangga Imam Gunawan, 2016.

John A. Dossey et al. (2006). *Discrete Mathematics*, fifth edition. Addition-Wesley. Pearson Education, Inc.

Joni Wilson Sitopu., (2023). *Bab 1 Konsep Teori Peluang*. Get Press Indonesia. Padang. Sumatera Barat.

Joni Wilson Sitopu., Nanang.,dkk.,(2023). *Teori Peluang*. Get Press Indonesia. Padang. Sumatera Barat.

Kapur, J.N dan Saxena, H.C. (1976)., *Mathematical Statistics*, eight Edition, S. Chand dan Campany LTD, New Delhi.

Rahmi Wahyuni, Suryanti.,dkk.,(2023). *Statistika*. Get Press Indonesia. Padang. Sumatera Barat.

Sheldon Ross,(2010). *A FIRST COURSE IN PROBABILITY*, Eighth Edition. University of Southern California. Pearson Education Upper Saddle River, New Jersey.

Sitopu. Joni Wilson, (1990). Fungsi Pembangkit momen dari beberapa distribusi probabilitas. *Prodi Matematika FMIPA USU* .Medan.

Sitopu, Joni Wilson., dkk, (2022). *Statistik Dasar*. PT. Global Eksekutif Teknologi. Padang.

Spiegel, Murray R, Ph.D.(2013)., Probability and Statistics. Schaum's outline series, Fourth Edition, McGRAW-HILL Book Company Net Work.

Walpole, R.E. 1992. *Pengantar Statistika.* Penerbit Gramedia.

Zul Fadli Rr., Suprantiningrum, Joni Wilson Sitopu.,dkk.,(2023). Statistik Ekonomi. PT. Global Eksekutif Teknologi. Padang.

# BAB 8
# METODE PENGAMBILAN SAMPEL

**Oleh Nadia Fazira**

## 8.1 Pendahuluan

Kualitas data menjadi salah satu komponen penting dalam menentukan keberhasilan sebuah penelitian baik dalam bidang ekonomi dan bisnis maupun bidang lainnya. Data yang digunakan haruslah akurat, relevan dan representatif sehingga temuan penelitian dapat dijadikan referensi dalam pengambilan keputusan, evaluasi kebijakan maupun pengembangan teori dan praktik ekonomi. Namun, mengumpulkan data dari seluruh populasi sering kali tidak memungkinkan untuk dilakukan karena membutuhkan biaya, waktu dan sumber daya yang tinggi. Oleh sebab itu, menggunakan sampel sebagai representasi dari populasi menjadi pilihan yang tepat dan efisien.

Dalam konteks penelitian bidang ekonomi dan bisnis, populasi merujuk pada seluruh kelompok atau unit yang menjadi objek pengamatan. Populasi ini dapat berupa individu (seperti konsumen, pelaku usaha, pekerja), institusi (seperti perusahaan, sekolah, koperasi), atau wilayah (seperti desa, kota, provinsi). Sementara itu, sampel adalah bagian dari populasi yang dipilih untuk dianalisis dengan harapan dapat memberikan gambaran umum dari keseluruhan populasi (Sugiyono, 2018). Misalnya, jika

peneliti ingin mengetahui pola konsumsi masyarakat di provinsi X, maka peneliti tidak harus mewawancarai seluruh penduduk di provinsi tersebut, melainkan cukup dengan mengambil sebagian dari penduduk tersebut dengan cara yang tepat dan mewakili karakteristik populasi. Begitu juga apabila perusahaan ingin mengetahui tingkat kepuasan pelanggan terhadap produk yang baru diluncurkan, maka perusahaan tidak perlu mewawancarai seluruh pelanggan, melainkan cukup dengan mengambil sebagian dari pelanggannya.

Selanjutnya, penting juga bagi peneliti untuk selalu menggunakan teknik sampling yang tepat dalam memilih sampel agar dapat merepresentasikan populasi secara akurat. Penggunaan teknik sampling yang tidak tepat dapat menghasilkan bias sehingga berdampak pada kesalahan dalam penarikan kesimpulan dan rekomendasi kebijakan. Hal ini pada akhirnya dapat mempengaruhi pengambilan keputusan dan perencanaan strategis (Sekaran and Bougie, 2016; Sugiyono, 2018). Oleh sebab itu, seorang peneliti haruslah merancang dengan cermat metode pengambilan sampel dengan selalu mempertimbangkan karakteristik populasi serta tujuan utama dari penelitian yang dilakukan. Peneliti juga dituntut untuk memahami prinsip-prinsip dasar dari masing-masing teknik agar mampu memilih metode yang tidak hanya praktis tetapi juga menjamin validitas dan reliabilitas data yang dikumpulkan.

Bab ini akan mengulas secara komprehensif mengenai konsep dasar pengambilan sampel, tujuan dan prinsip yang mendasarinya, klasifikasi teknik sampling yang umum digunakan serta bagaimana menentukan ukuran sampel yang tepat berdasarkan pertimbangan statistik. Selain itu, akan dibahas pula berbagai bentuk kesalahan yang dapat terjadi dalam proses sampling, baik *sampling error* maupun

*non-sampling error* serta cara meminimalkannya. Untuk memperkuat pemahaman, bab ini juga akan menyertakan contoh aplikasi teknik sampling dalam berbagai studi empiris di bidang ekonomi dan bisnis.

## 8.2 Tujuan dan Prinsip Pengambilan Sampel

Pengambilan sampel dilakukan tidak hanya untuk keperluan praktis, tetapi juga untuk memastikan bahwa data yang diperoleh memiliki kualitas yang dapat diandalkan. Berikut ini akan dibahas lebih lanjut mengenai tujuan dan prinsip dasar dalam pengambilan sampel yang wajib dipahami oleh setiap peneliti sebagai landasan dalam merancang studi yang kuat secara metodologis dan terpercaya.

### 8.2.1 Tujuan Pengambilan Sampel

1. Mewakili Populasi

   Salah satu tujuan utama dari pengambilan sampel dalam penelitian adalah untuk memperoleh gambaran yang dapat mewakili kondisi atau karakteristik populasi secara menyeluruh. Menurut Creswell (2014), penting bagi peneliti untuk memastikan bahwa sampel yang dipilih benar-benar mencerminkan populasi, agar hasil penelitian tidak hanya akurat secara statistik tetapi juga relevan untuk diterapkan dalam berbagai konteks kebijakan maupun bisnis. Misalnya, jika sebuah lembaga riset pasar ingin memahami preferensi konsumen terhadap produk makanan ringan di Indonesia, maka Lembaga tersebut tidak harus melakukan survei kepada seluruh penduduk Indonesia. Dengan memilih sekitar 1.000 responden dari berbagai latar belakang wilayah dan demografi, hasil yang diperoleh sudah dapat

digunakan untuk menggambarkan kecenderungan pasar secara nasional.

2. Efisiensi Biaya, Waktu dan Tenaga

   Menggunakan sampel memungkinkan peneliti untuk menghemat waktu, tenaga, dan biaya, tanpa harus mengorbankan kualitas data yang diperoleh. Ini menjadi sangat penting terutama ketika penelitian melibatkan populasi yang sangat besar atau tersebar di banyak lokasi. Sebagai contoh, dalam survei kepuasan pelanggan di sebuah perusahaan *e-commerce*, tidak perlu mewawancarai seluruh pengguna. Cukup dengan melibatkan sekitar 10% dari pengguna aktif selama tiga bulan terakhir, perusahaan sudah bisa mendapatkan gambaran yang cukup akurat dan bermanfaat untuk mengambil keputusan yang tepat.

3. Menghasilkan Data yang Objektif dan Dapat Diulang

   Menurut Neuman (2011), agar sebuah penelitian dianggap kredibel, teknik pengambilan sampelnya harus dirancang secara sistematis. Artinya, metode yang digunakan harus cukup jelas dan terstruktur sehingga bisa diikuti kembali oleh peneliti lain dan memungkinkan hasil penelitian diuji ulang dalam kondisi atau konteks yang serupa.

### 8.2.2 Prinsip-Prinsip Pengambilan Sampel

1. Representatif

   Sampel yang representatif harus mencerminkan karakteristik seluruh populasi, seperti usia, jenis kelamin, wilayah, tingkat pendidikan, atau status ekonomi. Jika hal ini diabaikan, maka hasil penelitian tidak dapat diterapkan secara umum kepada seluruh populasi. Misalnya, dalam penelitian perilaku keuangan rumah tangga, sangat penting untuk memastikan bahwa sampel

mencakup rumah tangga dengan berbagai kelompok pendapatan agar hasilnya tidak bias terhadap kelas ekonomi tertentu.

2. Objektivitas

   Proses pemilihan sampel harus bebas dari subjektivitas peneliti serta dilakukan secara acak. Teknik seperti *random sampling* digunakan untuk memastikan bahwa setiap anggota populasi memiliki peluang yang sama untuk dipilih sehingga mengurangi kemungkinan bias.

3. Dapat Diuji Ulang

   Agar sebuah penelitian dapat diterima secara ilmiah maka metode sampling yang digunakan harus memungkinkan replikasi dengan prosedur yang sama dan menghasilkan temuan yang serupa (Neuman, 2011). Oleh karena itu, teknik pengambilan sampel perlu disusun secara sistematis dan transparan agar dapat diterapkan kembali oleh peneliti lain untuk keperluan konfirmasi maupun evaluasi ulang.

4. Mengurangi Bias

   Bias dalam pengambilan sampel dapat menyebabkan hasil yang tidak akurat dan tidak dapat dipercaya. Bias pengambilan sampel terjadi ketika kelompok tertentu terlalu banyak atau terlalu sedikit terwakili dari populasi (Babbie, 2010). Sebagai contoh, dalam survei preferensi politik, jika hanya mengambil responden dari media sosial maka hasil yang diperoleh dapat bias karena tidak semua lapisan masyarakat aktif di platform digital.

5. Validitas dan Reliabilitas

   Validitas dalam pengambilan sampel berarti bahwa sampel yang diambil benar-benar mencerminkan karakteristik yang ingin diukur dalam populasi. Sementara itu, reliabilitas mengacu pada kemampuan pengukuran untuk memberikan hasil yang konsisten. Keduanya saling

melengkapi dalam penelitian, dimana sampel yang valid akan meningkatkan kemampuan untuk menggeneralisasi hasil, sementara sampel yang reliabel akan menghasilkan temuan yang konsisten.

6. Disesuaikan dengan Tujuan Penelitian

   Pemilihan metode sampling harus mempertimbangkan desain penelitian, jenis data yang dibutuhkan, serta konteks lapangan. Misalnya, dalam penelitian untuk mengukur prevalensi penyakit dalam populasi, teknik sampling yang digunakan harus dapat memastikan bahwa sampel mencakup individu dengan karakteristik yang relevan terhadap penyakit yang diteliti.

## 8.3 Klasifikasi Teknik Pengambilan Sampel

Teknik pengambilan sampel secara umum terbagi menjadi dua kategori utama, yaitu sampling probabilitas dan sampling non-probabilitas. Pemilihan teknik ini sangat bergantung pada tujuan penelitian, ketersediaan data, serta konteks lapangan.

### 8.3.1 Sampling Probabilitas

Sampling probabilitas adalah teknik pengambilan sampel di mana setiap elemen dalam populasi memiliki peluang yang sama dan dapat dihitung untuk dipilih menjadi sampel. Teknik ini dinilai sebagai metode paling ilmiah dan andal karena memungkinkan generalisasi hasil ke populasi secara luas (Sugiyono, 2018). Jenis-jenis sampling probabilitas sebagai berikut:

- *Simple Random Sampling*

  Dalam bukunya, Kerlinger (2006) mendefinisikan *simple random sampling* sebagai metode penarikan sampel dari populasi di mana

setiap anggota memiliki peluang yang sama untuk terpilih. Pemilihan dilakukan secara acak, misalnya dengan undian atau bantuan program komputer. Kelebihan dari metode ini adalah mampu menghindari bias dalam pemilihan sampel, karena tidak ada campur tangan subjektif dari peneliti. Selain itu, hasil yang diperoleh umumnya representatif sehingga memungkinkan generalisasi temuan ke populasi yang lebih luas. Namun, metode ini memiliki kekurangan dari sisi efisiensi terutama ketika populasi sangat besar dan tersebar luas karena dibutuhkan daftar populasi lengkap dan waktu pengambilan bisa menjadi lama. Contoh aplikasinya, sebuah perusahaan asuransi ingin mengukur tingkat kepuasan nasabah. Dari 10.000 nasabah, perusahaan secara acak memilih 1.000 orang untuk diwawancarai menggunakan generator angka acak.

- *Systematic Sampling*

  *Systematic sampling* dilakukan dengan memilih elemen dari populasi berdasarkan interval tertentu. Misalnya, jika ingin memilih 100 sampel dari populasi sebanyak 1.000 orang, maka peneliti bisa memilih setiap elemen ke-10 setelah menentukan titik awal secara acak. Kelebihan teknik ini adalah lebih sederhana dan cepat dibanding *simple random sampling*, serta tetap dapat -prinsip probabilistik. Namun, jika data populasi memiliki pola atau pengelompokan tertentu yang kebetulan sejalan dengan interval yang dipilih, maka hasilnya bisa menjadi bias. Sebagai contoh, dalam survei kualitas layanan pelanggan di sebuah perusahaan jasa logistik, pihak manajemen memilih pelanggan

ke-50, ke-100, dan seterusnya dari daftar pelanggan untuk dikontak melalui email.

- *Stratified Random Sampling*

  Teknik ini membagi populasi ke dalam beberapa strata atau kelompok homogen, lalu melakukan pengambilan sampel secara acak dari tiap kelompok tersebut. Stratifikasi bisa berdasarkan usia, jenis kelamin, wilayah geografis, atau status ekonomi (Hair et al., 2010). Metode ini sangat efektif dalam meningkatkan akurasi hasil karena tiap kelompok penting dalam populasi terwakili. Kelebihannya adalah mengurangi varians antar kelompok, terutama ketika karakteristik antar strata sangat berbeda. Kekurangannya, teknik ini membutuhkan data awal yang cukup mendetail, karena proses identifikasi dan pengelompokan strata tidak sederhana. Misalnya, dalam penelitian tentang penggunaan layanan *e-wallet*, peneliti membagi responden berdasarkan kelompok umur, lalu memilih sampel secara acak dari masing-masing kelompok tersebut agar hasilnya mencerminkan keseluruhan populasi pengguna e-wallet.

- *Cluster Sampling*

  *Cluster sampling* digunakan dengan cara membagi populasi menjadi beberapa kelompok atau klaster berdasarkan kriteria tertentu, biasanya geografis. Kemudian, beberapa klaster dipilih secara acak dan seluruh elemen dalam klaster terpilih menjadi sampel. Kelebihan metode ini terletak pada efisiensi biaya dan waktu karena peneliti cukup fokus pada wilayah-wilayah terpilih tanpa perlu mengakses seluruh populasi. Namun,

jika klaster-klaster tersebut tidak homogen atau terlalu bervariasi, hasil penelitian bisa memiliki varian besar dan kurang akurat dibanding teknik sampling lain. Contoh aplikatifnya dapat dilihat pada survei sosial ekonomi (SUSENAS) yang dilakukan BPS. BPS memilih beberapa desa atau kelurahan sebagai klaster lalu melakukan wawancara pada seluruh rumah tangga di wilayah tersebut.

### 8.3.2 Sampling Non-Probabilitas

Menurut Sugiyono (2018), sampling non-probabilitas adalah metode pengambilan sampel di mana tidak semua anggota populasi memiliki peluang yang sama untuk dipilih. Teknik ini sering digunakan dalam situasi keterbatasan akses data atau dalam penelitian eksploratif, dimana sampel dipilih berdasarkan pertimbangan tertentu, seperti kemudahan dijangkau atau relevansi informasi yang dimiliki. Jenis-jenis sampling non-probabilitas sebagai berikut:

1. *Convenience Sampling*

   *Convenience sampling* atau pengambilan sampel berdasarkan kemudahan akses yaitu dilakukan dengan memilih responden yang paling mudah dijangkau oleh peneliti. Teknik ini umumnya digunakan dalam penelitian eksploratif atau dalam situasi keterbatasan waktu dan sumber daya. Kelebihannya adalah praktis, murah, dan cepat namun sangat rentan terhadap bias karena tidak semua anggota populasi memiliki peluang yang sama untuk terpilih. Representativitas data pun rendah sehingga tidak bisa digeneralisasikan. Contoh aplikasinya, mahasiswa melakukan

penelitian perilaku konsumsi dan mengambil data dari teman-temannya di kampus yang sedang berada di kantin.

2. *Purposive Sampling*

   *Purposive sampling* adalah teknik pengambilan sampel dimana peneliti memilih subjek yang memiliki karakteristik khusus yang dianggap penting untuk memahami fenomena yang sedang diteliti (Creswell, 2014). Teknik ini sering digunakan dalam penelitian kualitatif atau ketika hanya kelompok tertentu yang memiliki informasi relevan. Kelebihannya, metode ini memungkinkan peneliti fokus pada responden yang tepat untuk menjawab pertanyaan penelitian secara mendalam. Namun, terdapat risiko subjektivitas tinggi dan hasil penelitian tidak dapat digeneralisasikan ke populasi luas. Sebagai contoh, peneliti ingin mengkaji adaptasi digital UMKM di era pandemi, lalu memilih secara khusus pelaku UMKM digital aktif di platform Tokopedia dan Shopee untuk diwawancarai.

3. *Quota Sampling*

   *Quota sampling* melibatkan penetapan kuota untuk masing-masing kategori responden, seperti kuota berdasarkan jenis kelamin, usia, atau pendapatan. Setelah kuota ditetapkan, peneliti mengisi kuota tersebut dengan responden sesuai karakteristik, meski tanpa proses randomisasi. Kelebihannya, teknik ini lebih cepat dan efisien dari stratified sampling namun tetap mengakomodasi keragaman dalam populasi. Kekurangannya, karena tidak ada unsur acak, tetap ada potensi bias pemilihan. Misalnya, dalam survei opini publik

tentang inflasi, peneliti menetapkan bahwa 50% sampel berasal dari kota dan 50% dari desa, lalu mencari responden sampai kuota tersebut terpenuhi.

4. *Snowball Sampling*

   Snowball sampling digunakan ketika populasi yang diteliti sulit diidentifikasi atau tergolong tersembunyi. Menurut Sugiyono (2018), jumlah sampel snowball sampling mula-mula kecil, kemudian sampel ini memilih teman-temannya untuk dijadikan sampel, begitu seterusnya, sehingga jumlah sampel tersebut menjadi banyak, seperti bola salju yang menggelinding makin lama makin besar. Kelebihannya adalah memungkinkan akses terhadap populasi kecil atau sulit dijangkau, namun metode ini tidak dapat menjamin representativitas dan memiliki potensi bias karena sangat tergantung pada jejaring sosial responden awal. Contoh aplikatifnya dapat dilihat pada penelitian terhadap pengguna kripto di pedesaan atau pelaku ekonomi informal tanpa registrasi resmi.

**Tabel 8. 1 Perbandingan Sampling Probabilitas dan Non-Probabilitas**

| Kriteria | Sampling Probabilitas | Sampling Non-Probabilitas |
|---|---|---|
| Representatif | Tingkat representatif tinggi karena setiap elemen populasi memiliki peluang yang sama untuk terpilih. | Representatif rendah hingga sedang karena pemilihan sampel tidak acak. |
| Kemungkinan Bias | Rendah, karena proses seleksi dilakukan secara acak dan sistematis. | Tinggi, karena bergantung pada subjektivitas atau aksesibilitas sampel. |
| Generalisasi | Hasil dapat digeneralisasikan ke seluruh populasi dengan lebih percaya diri. | Terbatas, karena tidak mencerminkan keseluruhan populasi secara utuh. |
| Kebutuhan Data Awal | Membutuhkan data populasi yang lengkap dan terstruktur. | Tidak selalu memerlukan data populasi lengkap; bisa dilakukan dengan akses terbatas. |
| Biaya dan Waktu | Relatif lebih tinggi karena proses seleksi yang sistematis dan menyeluruh. | Lebih efisien dari segi biaya dan waktu, cocok untuk penelitian eksploratif. |

Sumber: diolah oleh penulis

## 8.4 Menentukan Ukuran Sampel

Menentukan ukuran sampel merupakan langkah penting dalam proses pengambilan data. Sampel yang terlalu kecil berisiko menghasilkan hasil yang tidak akurat, sedangkan sampel yang terlalu besar bisa memboroskan sumber daya. Oleh karena itu, peneliti perlu menyesuaikan ukuran sampel dengan karakteristik populasi, tingkat kepercayaan dan *margin of error* yang diinginkan.

### 8.4.1 Rumus Penentuan Ukuran Sampel

Beberapa rumus yang sering digunakan dalam penelitian sosial dan ekonomi antara lain:

1. Rumus Slovin

   Digunakan ketika jumlah populasi diketahui dan peneliti ingin menentukan ukuran sampel dengan tingkat kesalahan tertentu.

   $$n = \frac{N}{1 + Ne^2}$$

   Keterangan:

   n : ukuran sampel

   N : jumlah populasi

   e: margin of error (misal: 0,05 untuk 5%)

2. Rumus Yamane

   Secara teknis, ini identik dengan Slovin, hanya berbeda dalam penyebutan referensi. Digunakan untuk populasi yang diketahui secara pasti.

3. Rumus Cochran

   Cocok digunakan untuk populasi besar (atau tak terhingga), terutama dalam survei berskala nasional atau studi akademik besar.

   $$n = \frac{Z^2.p.q}{e^2}$$

Keterangan:

n : ukuran sampel

Z : skor Z sesuai tingkat kepercayaan (misal: 1,96 untuk 95%)

p : proporsi responden (misal: 0,5 jika tidak diketahui)

q : 1-p

e: margin of error

### 8.4.2 Pertimbangan dalam Menentukan Ukuran Sampel

Beberapa hal yang perlu diperhatikan oleh peneliti sebelum menetapkan jumlah sampel:

1. Heterogenitas populasi

   Semakin beragam karakteristik dalam populasi maka semakin besar ukuran sampel yang dibutuhkan untuk menangkap variasi tersebut. Hal ini penting untuk memastikan bahwa sampel dapat merepresentasikan seluruh populasi dengan akurat.

2. Tingkat kepercayaan dan margin of error

   Tingkat kepercayaan yang lebih tinggi dan margin of error yang lebih kecil memerlukan ukuran sampel yang lebih besar untuk memastikan bahwa hasil penelitian mendekati nilai sebenarnya dalam populasi.

   ✓ Tujuan penelitian dan metode analisis

   Penelitian kuantitatif atau yang menggunakan analisis statistik inferensial biasanya memerlukan sampel yang lebih besar dibandingkan dengan penelitian eksploratif atau kualitatif. Hal ini untuk memastikan bahwa analisis statistik memiliki kekuatan yang cukup untuk mendeteksi efek atau perbedaan yang signifikan.

- ✓ Ketersediaan waktu dan dana

  Sumber daya yang tersedia, seperti waktu dan dana juga mempengaruhi ukuran sampel. Peneliti perlu menyeimbangkan kebutuhan metodologis dengan keterbatasan dilapangan.

### 8.4.3 Studi Kasus Aplikatif

Misalnya, sebuah UMKM di kota X ingin mengukur tingkat kepuasan pelanggan terhadap layanan *online*-nya. Total pelanggan aktif yang tercatat sebanyak 1.500 orang. Dengan *margin of error* sebesar 5%, maka peneliti menggunakan rumus Slovin untuk menghitung jumlah sampel:

$$n = \frac{N}{1 + Ne^2} = \frac{1500}{1 + 1500(0{,}05)^2} = \frac{1500}{1 + 3{,}75} \approx 316$$

Artinya, survei cukup dilakukan terhadap 316 pelanggan saja untuk memperoleh hasil yang valid dan efisien.

## 8.5 *Sampling Error* dan *Non-Sampling Error*

Dalam setiap proses pengambilan data melalui sampel, terdapat kemungkinan kesalahan yang memengaruhi akurasi hasil penelitian. Kesalahan ini umumnya dibagi menjadi dua jenis utama, yaitu *sampling error* dan *non-sampling error*.

### 8.5.1 *Sampling Error*

*Sampling error* terjadi ketika sampel yang diambil tidak sepenuhnya mewakili populasi, sehingga estimasi yang dihasilkan dari sampel bisa berbeda dari parameter sebenarnya jika seluruh populasi diteliti. Kesalahan ini merupakan bagian dari *total survey error*, yaitu keseluruhan kesalahan yang dapat memengaruhi

hasil survei (Biemer & Lyberg, 2003). Penyebab umumnya:

1. Ukuran sampel terlalu kecil
2. Teknik sampling yang kurang tepat
3. Variabilitas yang tinggi dalam populasi

Contohnya jika peneliti hanya mengambil responden dari wilayah perkotaan, padahal populasi juga mencakup wilayah pedesaan dengan karakteristik berbeda, maka hasilnya bisa bias terhadap kelompok wilayah perkotaan. Maka untuk meminimalkan bias tersebut dapat dilakukan beberapa cara berikut:

1. Gunakan teknik sampling yang tepat seperti *stratified* sampling untuk populasi yang heterogen
2. Pilih ukuran sampel yang memadai berdasarkan rumus dan pertimbangan statistik.

### 8.5.2 *Non-Sampling Error*

*Non-sampling error* adalah kesalahan yang muncul bukan karena pemilihan sampel, melainkan akibat dari berbagai faktor lain seperti kesalahan pengukuran, non-respons atau kesalahan dalam proses pengolahan data (Groves et al., 2009). Penyebab umumnya:

1. Kesalahan pengukuran misalnya responden salah mengisi kuesioner atau interviewer salah mencatat jawaban
   - ✓ Non-response error misalnya sebagian responden tidak menjawab atau menolak berpartisipasi
   - ✓ Kesalahan pemrosesan data seperti kesalahan saat input atau analisis data

- ✓ Desain instrumen yang buruk seperti pertanyaan tidak jelas, ambigu, atau menyesatkan.

Contohnya dalam survei perilaku konsumsi, jika responden merasa malu menjawab jujur terkait pengeluaran untuk rokok atau alkohol, maka data yang dikumpulkan bisa tidak valid. Maka untuk meminimalkan kesalahan tersebut dapat dilakukan beberapa cara berikut:

1. Uji coba kuesioner (pre-test) sebelum disebar
   - ✓ Latih enumerator agar tidak mengarahkan jawaban responden
   - ✓ Gunakan bahasa yang netral dan mudah dipahami dalam instrumen.

### 8.5.3 Implikasi Terhadap Hasil Penelitian

Kedua jenis *error* ini dapat menurunkan validitas dan reliabilitas hasil penelitian. Jika tidak dikendalikan dengan baik maka akan menyebabkan:

1. Kesimpulan menjadi keliru
2. Kebijakan berbasis data menjadi tidak efektif
3. Generalisasi terhadap populasi menjadi tidak dapat dipertanggungjawabkan.

Oleh karena itu, peneliti harus secara aktif merancang strategi untuk mengidentifikasi, mengukur dan meminimalkan *error* sejak tahap awal penelitian.

## 8.6 Contoh Aplikasi dalam Ekonomi dan Bisnis

Penggunaan teknik pengambilan sampel dalam penelitian ekonomi dan bisnis sangat penting untuk menghasilkan data yang representatif dan dapat diandalkan. Beberapa studi empiris telah menggunakan berbagai metode sampling untuk menganalisis fenomena ekonomi, sosial dan bisnis. Berikut adalah beberapa contoh aplikasi metode pengambilan sampel dalam penelitian di bidang ekonomi dan bisnis:

1. Survei Kepuasan Pelanggan E-Commerce

   Dalam penelitian "Pengaruh Kualitas Layanan terhadap Kepuasan Pelanggan Tokopedia" oleh Sari dan Prasetyo (2020), digunakan teknik purposive sampling untuk memilih responden yang merupakan pengguna aktif Tokopedia. Teknik ini dipilih karena peneliti membutuhkan responden dengan karakteristik tertentu seperti frekuensi transaksi dan lama penggunaan aplikasi. Meskipun bukan teknik probabilistik, purposive sampling dalam penelitian ini efektif untuk mengungkap preferensi dan kepuasan pengguna secara mendalam.

2. Preferensi Konsumen terhadap Produk Lokal

   Dalam artikel "Analisis Preferensi Konsumen Terhadap Produk Makanan Lokal di Yogyakarta" oleh Andriani dan Kurniawan (2021), digunakan *cluster sampling* untuk menjangkau konsumen dari berbagai kecamatan di wilayah tersebut. Setiap klaster dipilih berdasarkan kedekatan geografis dan responden dipilih secara acak di dalamnya. Pendekatan ini sangat berguna untuk mengatasi kendala geografis dan efisiensi biaya dalam pengumpulan data.

## 8.7 Penutup

Pengambilan sampel merupakan komponen yang krusial dalam proses penelitian terutama dalam bidang ekonomi dan bisnis yang sering kali berhadapan dengan populasi yang besar dan heterogen. Melalui teknik sampling yang tepat maka peneliti dapat memperoleh data yang akurat, efisien dan representatif tanpa harus mengobservasi seluruh populasi. Prinsip-prinsip seperti representativitas, objektivitas, replikasi dan pengurangan bias menjadi dasar dalam merancang strategi sampling yang andal.

Terdapat dua klasifikasi utama teknik sampling yaitu sampling probabilitas dan non-probabilitas, dimana keduanya memiliki keunggulan dan keterbatasan tergantung pada tujuan dan konteks penelitian. Pemilihan metode sampling harus mempertimbangkan faktor-faktor seperti struktur populasi, ketersediaan data, sumber daya yang dimiliki, serta tingkat akurasi dan generalisasi yang diharapkan dari hasil penelitian.

Selain itu, penentuan ukuran sampel juga menjadi aspek penting yang harus dirancang secara sistematis dengan mempertimbangkan *margin of error*, tingkat kepercayaan serta heterogenitas populasi. Teknik-teknik seperti rumus Slovin, Cochran, dan Yamane dapat digunakan sebagai acuan untuk menentukan jumlah sampel yang tepat. Pemahaman terhadap potensi *sampling error* dan *non-sampling error* juga tidak kalah penting untuk menjaga kualitas dan validitas data yang diperoleh. Kesadaran akan potensi kesalahan ini akan membantu peneliti merancang strategi mitigasi yang sesuai.

Berbagai contoh empiris telah menunjukkan bagaimana teknik sampling digunakan dalam praktik. Ini menegaskan bahwa pemilihan metode pengambilan sampel yang tepat bukan hanya persoalan teknis tetapi juga strategis dalam

menghasilkan pengetahuan yang bermanfaat bagi pengembangan teori dan kebijakan.

Dengan memahami konsep, prinsip dan teknik sampling secara komprehensif, peneliti diharapkan mampu merancang proses pengumpulan data yang bermutu tinggi serta mampu menjawab pertanyaan penelitian secara meyakinkan dan akurat.

## DAFTAR PUSTAKA

Andriani, D. and Kurniawan, F., 2021. *Analisis Preferensi Konsumen Terhadap Produk Makanan Lokal di Yogyakarta*. *Jurnal Ekonomi dan Pembangunan Daerah*.

Babbie, E., 2010. *The Practice of Social Research*. 12th ed. Belmont, CA: Wadsworth.

Biemer, P.P. and Lyberg, L.E., 2003. *Introduction to Survey Quality*. Hoboken, NJ: Wiley.

Creswell, J.W., 2014. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 4th ed. Thousand Oaks, CA: SAGE Publications.

Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E. and Tourangeau, R., 2009. *Survey Methodology*. 2nd ed. Hoboken, NJ: Wiley.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E., 2010. *Multivariate Data Analysis*. 7th ed. Upper Saddle River, NJ: Pearson Prentice Hall.

Kerlinger, F.N., 2006. *Foundations of Behavioral Research*. 4th ed. New York: Holt, Rinehart and Winston.

Neuman, W.L., 2011. *Social Research Methods: Qualitative and Quantitative Approaches*. 7th ed. Boston: Pearson Education.

Sari, R. and Prasetyo, A., 2020. *Pengaruh Kualitas Layanan terhadap Kepuasan Pelanggan Tokopedia*. *Jurnal Ekonomi dan Bisnis Digital Indonesia*.

Sekaran, U. and Bougie, R., 2016. *Research Methods for Business: A Skill-Building Approach*. 7th ed. Chichester: John Wiley & Sons.

Sugiyono, 2018. *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.

# BAB 9
# UJI HIPOTESIS

**Oleh Asyidatur Rosmaniar**

## 9.1 Pendahuluan

Dalam dunia bisnis dan ekonomi, pengambilan keputusan yang tepat memerlukan dasar yang kuat, terutama ketika berhadapan dengan ketidakpastian. Banyak keputusan penting, seperti strategi pemasaran, kebijakan harga, atau bahkan keputusan investasi, sering kali bergantung pada data dan analisis statistik. Salah satu metode yang paling umum digunakan untuk mendukung keputusan ini adalah uji hipotesis. Uji hipotesis adalah teknik statistik yang digunakan untuk menguji klaim atau dugaan mengenai parameter populasi berdasarkan sampel data yang diperoleh (Lind, Marchal and Wathen, 2012; Anderson *et al.*, 2014).

Secara sederhana, uji hipotesis memungkinkan peneliti atau pengambil keputusan untuk memverifikasi atau menolak suatu dugaan (hipotesis) tentang karakteristik populasi (Sugiyono, 2014; Riduwan, 2020). Misalnya, seorang manajer mungkin ingin mengetahui apakah rata-rata penjualan produk di satu toko lebih tinggi dibandingkan dengan toko lain. Atau, apakah kebijakan baru dalam suatu perusahaan memiliki efek yang signifikan terhadap kepuasan pelanggan. Dengan menggunakan uji hipotesis, mereka dapat membuat keputusan yang lebih objektif berdasarkan bukti yang ada (Anderson *et al.*, 2014).

Uji hipotesis sangat penting dalam ekonomi dan bisnis karena banyak fenomena yang tidak dapat langsung diamati atau diukur (Hurriyati and Gunarto, 2019). Oleh karena itu, kita sering menggunakan sampel dari populasi yang lebih besar untuk membuat inferensi atau kesimpulan tentang populasi tersebut. Di sini, uji hipotesis berfungsi untuk menentukan apakah data yang terkumpul cukup kuat untuk mendukung klaim yang dibuat tentang populasi (Lind, Marchal and Wathen, 2012; Nica, Dean and Illowsky, 2013).

Pada bab ini, kita akan mempelajari dasar-dasar uji hipotesis, langkah-langkah dalam melakukan uji, serta bagaimana uji hipotesis dapat diterapkan dalam konteks ekonomi dan bisnis. Dengan pemahaman yang mendalam tentang uji hipotesis, para profesional di bidang bisnis dan ekonomi dapat membuat keputusan yang lebih tepat dan berbasis data, yang pada akhirnya meningkatkan kinerja dan efisiensi organisasi.

## 9.2 Konsep Dasar Uji Hipotesis

Uji hipotesis adalah suatu metode statistik yang digunakan untuk menguji suatu klaim atau dugaan tentang parameter populasi menggunakan data sampel. Dalam penerapannya, uji hipotesis melibatkan beberapa konsep dasar yang penting untuk dipahami.

### 9.2.1 Hipotesis Statistik

- **Hipotesis Nol ($H_0$)**

  Hipotesis nol, atau $H_0$, merupakan pernyataan yang menyatakan bahwa tidak ada perbedaan, hubungan, atau efek yang signifikan dalam populasi yang sedang diuji. Sebagai contoh, dalam uji perbandingan rata-rata dua kelompok, hipotesis nol bisa berbunyi "rata-rata

kedua kelompok adalah sama." Hipotesis nol ini sering kali dijadikan dasar untuk pengujian, karena uji hipotesis umumnya berfungsi untuk menguji apakah data yang ada cukup kuat untuk menolak $H_0$.

- **Hipotesis Alternatif ($H_1$ atau Ha)**

Hipotesis alternatif, yang dilambangkan dengan $H_1$ atau Ha, merupakan klaim yang bertentangan dengan hipotesis nol. Ini menyatakan bahwa terdapat perbedaan atau efek yang signifikan antara dua kelompok atau variabel. Misalnya, hipotesis alternatif dalam uji perbandingan rata-rata dua kelompok dapat berbunyi "rata-rata kedua kelompok tidak sama." Uji hipotesis bertujuan untuk mencari bukti yang cukup untuk mendukung hipotesis alternatif ini.

### 9.2.2 Jenis Kesalahan dalam Pengujian Hipotesis

Dalam uji hipotesis, ada dua jenis kesalahan yang dapat terjadi, yaitu kesalahan tipe I dan kesalahan tipe II.

- **Kesalahan Tipe I (α)**

Kesalahan tipe I terjadi ketika hipotesis nol ditolak padahal sebenarnya benar. Ini berarti kita salah mengklaim bahwa ada perbedaan atau hubungan yang signifikan, padahal kenyataannya tidak ada. Probabilitas kesalahan tipe I dinyatakan dengan level signifikansi (α), yang seringkali ditetapkan sebesar 0,05 (5%).

- **Kesalahan Tipe II (β)**

Kesalahan tipe II terjadi ketika kita gagal menolak hipotesis nol padahal sebenarnya hipotesis alternatif yang benar. Ini berarti kita tidak dapat mendeteksi efek yang seharusnya ada dalam data. Probabilitas kesalahan tipe II disebut sebagai β, dan 1 - β adalah kekuatan uji

(*power of the test*), yang menunjukkan kemampuan uji untuk mendeteksi perbedaan yang benar-benar ada.

### 9.2.3 Level Signifikansi dan Kekuatan Uji (P*ower of the Test*)

- **Level Signifikansi (α)**

Level signifikansi (α) adalah probabilitas maksimum yang diterima untuk membuat kesalahan tipe I, yaitu menolak hipotesis nol yang sebenarnya benar. Nilai α yang umum digunakan dalam banyak penelitian adalah 0,05, yang berarti kita menerima risiko 5% untuk membuat kesalahan tipe I dalam pengujian.

- **Kekuatan Uji (*Power of the Test*)**

Kekuatan uji (*power of the test*) adalah kemampuan uji untuk menolak hipotesis nol ketika hipotesis alternatif benar, yaitu kemampuannya untuk meng-hindari kesalahan tipe II. Semakin tinggi kekuatan uji, semakin besar kemungkinan untuk mendeteksi perbedaan atau efek yang benar-benar ada.

### 9.2.4 Konsep Nilai Kritis dan *p-value*

- **Nilai Kritis**

Nilai kritis adalah nilai batas yang digunakan untuk memutuskan apakah statistik uji yang dihitung berada dalam daerah penolakan atau tidak. Jika statistik uji melebihi nilai kritis, maka kita menolak hipotesis nol. Nilai kritis ditentukan berdasarkan level signifikansi α yang telah ditentukan sebelumnya dan jenis uji yang digunakan.

- ***p-value***

p-value adalah probabilitas untuk mendapatkan hasil yang lebih ekstrem atau lebih jauh dari nilai yang diamati dalam data, dengan asumsi hipotesis nol benar. Jika p-value lebih kecil dari level signifikansi α, kita menolak hipotesis nol. Sebaliknya, jika p-value lebih besar dari α, kita gagal menolak hipotesis nol. p-value memberikan informasi yang lebih spesifik tentang seberapa kuat bukti yang ada terhadap hipotesis nol.

## 9.3 Langkah-Langkah dalam Uji Hipotesis

Proses uji hipotesis melibatkan beberapa langkah yang harus diikuti dengan cermat untuk mencapai keputusan yang tepat berdasarkan data yang tersedia. Berikut adalah langkah-langkah dalam melakukan uji hipotesis:

### 9.3.1 Menentukan Hipotesis ($H_0$ dan $H_1$)

Langkah pertama dalam uji hipotesis adalah merumuskan dua hipotesis yang saling bertentangan, yaitu hipotesis nol ($H_0$) dan hipotesis alternatif ($H_1$ atau Ha). Hipotesis nol ($H_0$) merupakan pernyataan bahwa tidak ada perbedaan atau efek yang signifikan dalam populasi yang diuji. Sebaliknya, hipotesis alternatif ($H_1$) adalah pernyataan yang menyatakan bahwa ada perbedaan atau efek yang signifikan. Misalkan seorang manajer pemasaran ingin mengetahui apakah ada perbedaan rata-rata penjualan antara dua cabang toko. Hipotesis yang dirumuskan adalah:

- Hipotesis Nol ($H_0$): "Rata-rata penjualan cabang A sama dengan rata-rata penjualan cabang B."

$$H_0: \mu A = \mu B$$

- Hipotesis Alternatif ($H_1$): "Rata-rata penjualan cabang A tidak sama dengan rata-rata penjualan cabang B."

$$H_1: \mu A \neq \mu B H_1$$

Pada kasus ini, kita ingin menguji apakah ada perbedaan rata-rata penjualan antara kedua cabang.

### 9.3.2 Menentukan Level Signifikansi (α)

Setelah merumuskan hipotesis, langkah berikutnya adalah menentukan level signifikansi (α). Level signifikansi adalah probabilitas maksimum yang diterima untuk melakukan kesalahan tipe I, yaitu menolak hipotesis nol yang sebenarnya benar. Biasanya, level signifikansi ditetapkan sebesar 0,05 (5%), yang berarti kita siap menerima kemungkinan 5% untuk membuat kesalahan tipe I. Level signifikansi ini akan digunakan untuk menentukan batasan nilai statistik uji yang dapat diterima.

### 9.3.3 Memilih Statistik Uji yang Tepat

Pemilihan statistik uji yang tepat sangat penting untuk melakukan uji hipotesis yang valid. Statistik uji yang digunakan bergantung pada jenis data dan hipotesis yang diuji. Beberapa jenis statistik uji yang sering digunakan meliputi:

- Uji t (untuk rata-rata satu sampel atau dua sampel ketika varians tidak diketahui),
- Uji z (untuk rata-rata satu sampel atau dua sampel ketika varians diketahui),
- Uji *chi-square* (untuk data kategori atau uji kecocokan),

- Uji F (untuk analisis varians atau perbandingan lebih dari dua kelompok)

Misalnya, untuk uji z ($\sigma$ diketahui) nilai statistik uji dapat dihitung dengan rumus:

$$z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

di mana:

- $\overline{X}$ adalah rata-rata sampel,
- $\mu_0$ adalah rata-rata populasi yang diuji,
- $\sigma$ adalah deviasi standar sampel,
- n adalah ukuran sampel

### 9.3.4 Menentukan Aturan Keputusan (Daerah Kritis)

Aturan keputusan ditentukan berdasarkan nilai kritis yang diperoleh dari distribusi statistik uji. Daerah kritis adalah area di luar nilai kritis, di mana jika statistik uji yang dihitung berada dalam daerah ini, maka hipotesis nol akan ditolak. Nilai kritis ini ditentukan oleh level signifikansi ($\alpha$) yang telah ditetapkan sebelumnya, dan jenis uji yang dilakukan (satu sisi atau dua sisi).

### 9.3.5 Mengambil Keputusan

Setelah menghitung nilai statistik uji, langkah selanjutnya adalah membandingkan nilai statistik uji dengan nilai kritis. Jika nilai statistik uji jatuh dalam daerah kritis, maka hipotesis nol ditolak. Sebaliknya, jika nilai statistik uji berada di luar daerah kritis, hipotesis nol tidak ditolak. Keputusan ini bergantung pada apakah

p-value lebih kecil dari level signifikansi α atau tidak. Jika *p-value* < α, maka kita menolak hipotesis nol.

Contoh : seorang manajer pemasaran ingin mengetahui apakah ada perbedaan rata-rata penjualan antara dua cabang toko.

1. Menentukan hipotesis

   Hipotesis yang dirumuskan adalah:

   - Hipotesis Nol ($H_0$): "Rata-rata penjualan cabang A sama dengan rata-rata penjualan cabang B."

     $$H_0:\mu A=\mu BH_0$$

   - Hipotesis Alternatif ($H_1$): "Rata-rata penjualan cabang A tidak sama dengan rata-rata penjualan cabang B."

     $$H1:\mu A\neq\mu BH_1$$

   Pada kasus ini, kita ingin menguji apakah ada perbedaan rata-rata penjualan antara kedua cabang.

2. Menentukan Level Signifikansi (α)

   Manajer memilih level signifikansi α = 0,05 (5%), yang berarti bahwa mereka bersedia menerima risiko 5% untuk membuat kesalahan tipe I, yaitu menolak hipotesis nol padahal sebenarnya hipotesis nol benar.

3. Memilih Statistik Uji yang Tepat

   Karena kita membandingkan rata-rata dua kelompok (penjualan cabang A dan B), dan ukuran sampelnya besar, kita akan menggunakan Uji t dua sampel independen. Statistik uji yang digunakan adalah:

$$t = \frac{\overline{X_A} - \overline{X_B}}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

Di mana:

- $\overline{X_A}, \overline{X_B}$ adalah rata-rata sampel untuk cabang A dan B,
- $s_A^2, s_B^2$ adalah deviasi standar sampel untuk cabang A dan B,
- $n_A, n_B$ adalah ukuran sampel untuk cabang A dan B.

4. Menentukan Aturan Keputusan (Daerah Kritis)

   Dengan level signifikansi α = 0,05, kita menggunakan distribusi t untuk menentukan nilai kritis. Karena uji ini dua sisi, daerah kritisnya adalah di dua ujung distribusi t, masing-masing dengan probabilitas 0,025 di setiap sisi. Jika nilai statistik uji t yang dihitung lebih besar dari nilai t kritis atau lebih kecil dari -t kritis, maka kita akan menolak hipotesis nol.

5. Mengambil keputusan dan menginterpretasikan hasil

   Setelah mengumpulkan data penjualan dari kedua cabang, kita mendapatkan rata-rata penjualan berikut:

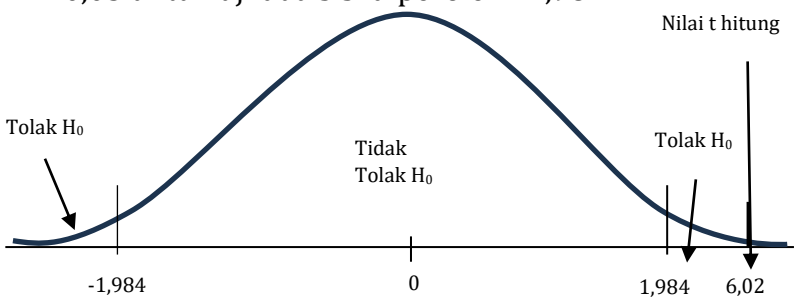
   - Rata-rata penjualan cabang A ( $\overline{X_A}$ ) = 120 unit
   - Rata-rata penjualan cabang B ($\overline{X_B}$ ) = 100 unit

- Deviasi standar cabang A ($s_A$) = 15
- Deviasi standar cabang B ($s_B$) = 18
- Ukuran sampel cabang A ($n_A$) = 50
- Ukuran sampel cabang B ($n_B$) = 50

Menggunakan rumus t, kita menghitung statistik uji t:

$$t = \frac{120 - 100}{\sqrt{\frac{15^2}{50} + \frac{18^2}{50}}} = \frac{20}{\sqrt{10{,}98}} \approx 6{,}02$$

Dari tabel distribusi t dengan df = 98 (50 + 50 - 2), kita mencari nilai kritis t pada level signifikansi 0,05 untuk uji dua sisi diperoleh ±1,984.




Karena nilai statistik uji t kita (6,02) lebih besar dari nilai kritis t (1,984), maka kita menolak hipotesis nol. Ini berarti ada bukti yang cukup untuk menyimpulkan bahwa rata-rata penjualan antara kedua cabang toko tersebut berbeda secara signifikan pada tingkat signifikansi 5%. Manajer

pemasaran dapat menggunakan informasi ini untuk mengambil keputusan lebih lanjut mengenai strategi pemasaran di kedua cabang.

## 9.4 Jenis-Jenis Uji Hipotesis dalam Statistik Ekonomi dan Bisnis

Dalam statistik ekonomi dan bisnis, berbagai jenis uji hipotesis digunakan untuk menguji berbagai klaim atau dugaan mengenai populasi berdasarkan data sampel. Berikut adalah beberapa jenis uji hipotesis yang sering digunakan:

### 9.4.1 Uji Hipotesis Satu Sampel (*One-Sample Test*)

Uji hipotesis satu sampel digunakan untuk menguji apakah rata-rata atau proporsi suatu populasi berbeda dari nilai tertentu yang diketahui atau diasumsikan.

**Uji Rata-rata Satu Sampel**

Uji rata-rata satu sampel digunakan untuk menguji apakah rata-rata suatu sampel berbeda signifikan dari nilai yang diasumsikan atau diketahui. Misalnya, perusahaan ingin mengetahui apakah rata-rata penjualan produk mereka berbeda dari target penjualan yang telah ditetapkan.

Persamaan untuk Uji t Satu Sampel:

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}}$$

di mana:

- $\overline{X}$ adalah rata-rata sampel,
- $\mu$ adalah rata-rata populasi yang diuji,
- $s$ adalah deviasi standar sampel,
- n adalah ukuran sampel

**Uji Proporsi Satu Sampel**

Uji proporsi satu sampel digunakan untuk menguji apakah proporsi suatu kategori dalam populasi berbeda dari nilai yang diketahui atau diasumsikan. Misalnya, kita ingin menguji apakah proporsi pelanggan yang puas dengan layanan perusahaan berbeda dari 80%.

Persamaan untuk Uji Proporsi Satu Sampel:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

Di mana:

- $\hat{p}$ = proporsi sampel
- $p_0$ = proporsi yang diharapkan dalam populasi
- $n$ = ukuran sampel

### 9.4.2 Uji Hipotesis Dua Sampel (*Two-Sample Test*)

Uji hipotesis dua sampel digunakan untuk membandingkan dua kelompok yang berbeda dan menguji apakah terdapat perbedaan yang signifikan antara kedua kelompok tersebut.

#### Uji Dua Sampel Independen (*Independent Samples Test*)

Uji dua sampel independen digunakan untuk menguji apakah rata-rata atau proporsi dua kelompok independen berbeda secara signifikan. Misalnya, kita ingin menguji apakah rata-rata pengeluaran antara dua kota berbeda.

Persamaan untuk Uji t Dua Sampel Independen:

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Di mana:

- $\overline{X_1}, \overline{X_2}$ adalah rata-rata sampel dua kelompok,
- $s_1^2, s_2^2$ adalah deviasi standar sampel dari dua kelompok,
- $n_1, n_2$ adalah ukuran sampel dari dua kelompok.

#### Uji Dua Sampel Berpasangan (*Paired Samples Test*)

Uji dua sampel berpasangan digunakan untuk menguji perbedaan antara dua kelompok yang saling berpasangan. Biasanya digunakan ketika pengukuran dilakukan pada subjek yang sama sebelum dan sesudah perlakuan.

Persamaan untuk Uji t Dua Sampel Berpasangan:

$$t = \frac{\bar{d}}{s \Big/ \sqrt{n}}$$

Di mana:

- $\bar{d}$ = rata-rata perbedaan antara pasangan data
- $s$ = deviasi standar perbedaan pasangan data
- $n$ = jumlah pasangan data

### 9.4.3 Uji Analisis Variansi (*Analysis of Variance* / ANOVA)

ANOVA digunakan untuk menguji apakah terdapat perbedaan rata-rata antara tiga kelompok atau lebih. Uji ini sangat berguna ketika kita ingin membandingkan lebih dari dua kelompok dan menguji apakah perbedaan rata-rata antar kelompok tersebut signifikan.

Persamaan untuk Uji ANOVA (Satu Arah):

$$F = \frac{variansi\ antar\ kelompok}{variansi\ dalam\ kelompok} = \frac{SSB \Big/ k-1}{SSW \Big/ n-1}$$

Di mana:

- $SSB$ = jumlah kuadrat antar kelompok
- $SSW$ = jumlah kuadrat dalam kelompok

- $k$ = jumlah kelompok
- $n$ = total ukuran sampel

Jika nilai F yang dihitung lebih besar dari nilai F kritis dari distribusi F, maka hipotesis nol ditolak.

### 9.4.4 Uji Chi-Square untuk Data Kategori

Uji chi-square digunakan untuk menguji hubungan antara dua variabel kategori. Uji ini berguna untuk menentukan apakah distribusi frekuensi yang diamati berbeda dari distribusi yang diharapkan. Uji chi-square juga digunakan dalam uji kebaikan fit dan uji independensi.

Persamaan untuk Uji Chi-Square:

$$x^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Di mana:

- $O_i$ = frekuensi observasi pada kategori ke-i
- $E_i$ = frekuensi yang diharapkan pada kategori ke-i
- $x^2$ adalah statistik uji chi-square

Jika nilai $x^2$ yang dihitung lebih besar dari nilai $x^2$ kritis, maka hipotesis nol ditolak, yang berarti ada hubungan signifikan antara variabel kategori yang diuji.

### 9.4.5 Uji Signifikansi dalam Analisis Regresi Linier

Dalam analisis regresi linier, uji signifikansi digunakan untuk menguji apakah koefisien regresi berbeda secara signifikan dari nol. Ini digunakan untuk menentukan apakah ada hubungan yang signifikan

antara variabel independen dan dependen dalam model regresi.

Persamaan untuk Uji t pada Koefisien Regresi:

$$t = \frac{b}{SE_b}$$

Di mana:

- $b$ = koefisien regresi (slope) yang diestimasi
- $SE_b$ = kesalahan standar dari koefisien regresi
- $t$ adalah statistik uji t untuk koefisien regresi

Jika nilai t yang dihitung lebih besar dari nilai t kritis, maka koefisien regresi dianggap signifikan pada level signifikansi yang ditentukan, dan kita dapat menyimpulkan bahwa variabel independen memiliki pengaruh yang signifikan terhadap variabel dependen.

## DAFTAR PUSTAKA

Anderson, D. R. *et al.* (2014) 'Statiscis for Business Economics', pp. 1–1204.

Hurriyati, R. and Gunarto, M. (2019) *Metode Statistika Bisnis untuk Bidang Ilmu Manajemen dengan Aplikasi Program SPSS*. Bandung: PT Refika Aditama.

Lind, D. A., Marchal, W. G. and Wathen, S. A. (2012) *Statistical Techniques in Bsiness & Economics 15th Edition*.

Nica, M., Dean, S. and Illowsky, B. P. D. (2013) 'Principles of Business Statistics', *Rice University, Houston, Texas*, p. 113. Available at: http://cnx.org/content/col10874/1.5/.

Riduwan (2020) *Dasar-Dasar Statistika*. Bandung: Penerbit Alfabeta.

Sugiyono (2014) *Statistika untuk Penelitian*. Bandung: Penerbit Alfabeta.

# BAB 10
# ANALISIS VARINS (ANOVA)

**Oleh Derry Nugraha**

## 10.1 Pengenalan ANOVA

### 10.1.1 Definisi dan Sejarah Singkat ANOVA

Analisis Varians (*Analysis of Variance* atau ANOVA) adalah teknik statistik yang digunakan untuk menganalisis perbedaan rata-rata antara tiga atau lebih kelompok dengan membandingkan variabilitas antara kelompok terhadap variabilitas dalam kelompok (Gut, 2009). ANOVA dikembangkan oleh statistikawan dan genetikawan Sir Ronald Fisher pada tahun 1920-an dan dipublikasikan pertama kali dalam bukunya "*Statistical Methods for Research Workers*" (1925).

Fisher mengembangkan ANOVA saat bekerja di *Rothamsted Experimental Station* di Inggris, sebuah pusat penelitian pertanian. Pengembangan ANOVA bermula dari kebutuhan untuk menganalisis hasil eksperimen pertanian yang kompleks dengan banyak variabel dan kelompok perlakuan (Dabas, 2024). Melalui kontribusinya yang signifikan, Fisher tidak hanya menciptakan ANOVA tetapi juga meletakkan dasar bagi pengembangan desain eksperimen modern.

Pada awalnya, ANOVA diterapkan terutama dalam bidang biologi dan pertanian, namun seiring waktu metode ini diadopsi secara luas dalam berbagai disiplin ilmu seperti psikologi, kedokteran, ekonomi, pendidikan, dan ilmu sosial. Perkembangan komputasi modern telah memudahkan penerapan ANOVA dalam analisis data kompleks, menjadikannya salah satu alat statistik yang paling sering digunakan dalam penelitian kuantitatif.

### 10.1.2 Tujuan dan Manfaat Penggunaan ANOVA dalam Analisis Statistik

ANOVA memiliki beberapa tujuan dan manfaat utama dalam analisis statistik (Brown & Kass, 2009):

a. Pengujian Perbedaan Kelompok
   ANOVA memungkinkan peneliti untuk menentukan apakah terdapat perbedaan yang signifikan antara rata-rata beberapa kelompok sekaligus, tanpa perlu melakukan multiple pairwise comparisons yang dapat meningkatkan risiko kesalahan Tipe I.
b. Efisiensi Statistik
   Dengan menggunakan satu pengujian untuk membandingkan beberapa kelompok, ANOVA lebih efisien dibandingkan melakukan serangkaian uji-t terpisah.
c. Pengendalian Error Rate
   ANOVA membantu mengontrol familywise error rate yang muncul ketika melakukan multiple comparisons.
d. Analisis Faktor Interaksi
   Dalam ANOVA multifaktor, peneliti dapat menganalisis tidak hanya efek utama dari setiap faktor tetapi juga interaksi antar faktor,

memberikan pemahaman yang lebih komprehensif tentang kompleksitas data.

e. Fleksibilitas Desain
   ANOVA dapat diterapkan pada berbagai desain penelitian, termasuk desain acak lengkap, desain blok, dan desain faktorial.
f. Partisi Variabilitas
   ANOVA memungkinkan peneliti untuk memisahkan dan mengkuantifikasi kontribusi berbagai sumber variasi dalam data, membantu mengidentifikasi faktor-faktor yang paling berpengaruh.
g. Dasar untuk Model Kompleks
   ANOVA menjadi fondasi untuk model statistik yang lebih kompleks seperti ANCOVA, MANOVA, dan model campuran (*mixed models*).

### 10.1.3 Perbedaan ANOVA dengan Uji-t

Meskipun ANOVA dan uji-t keduanya adalah metode untuk membandingkan rata-rata kelompok, terdapat beberapa perbedaan penting (Bulman & Osborn, 1989):

a. Jumlah Kelompok
   Uji-t dirancang untuk membandingkan dua kelompok saja, sedangkan ANOVA dapat membandingkan tiga atau lebih kelompok secara simultan.
b. Pendekatan Analisis
   Uji-t menguji perbedaan antara dua rata-rata dengan membandingkan t-statistik dengan distribusi-t kritis. ANOVA membandingkan variabilitas antara kelompok dengan variabilitas dalam kelompok, menghasilkan statistik F yang dibandingkan dengan distribusi F kritis.

c. Kontrol Kesalahan Tipe I
   Melakukan multiple uji-t untuk membandingkan banyak kelompok akan meningkatkan risiko kesalahan Tipe I (menolak $H_0$ yang benar). ANOVA mengatasi masalah ini dengan melakukan satu pengujian komprehensif.
d. Informasi Pasca Pengujian
   Uji-t memberikan informasi langsung tentang arah perbedaan antara dua kelompok. ANOVA hanya menunjukkan ada tidaknya perbedaan signifikan di antara kelompok, tanpa menentukan kelompok mana yang berbeda. Untuk menentukan perbedaan antar kelompok spesifik, diperlukan uji lanjut (*post-hoc tests*).
e. Power Statistik
   Dalam kondisi tertentu, ANOVA memiliki power statistik yang lebih besar dibandingkan serangkaian uji-t terpisah, terutama ketika ukuran sampel kecil atau efek yang dideteksi relatif kecil.
f. Asumsi
   Meskipun kedua metode memiliki asumsi dasar yang serupa (normalitas, homogenitas varians), ANOVA memiliki asumsi tambahan terkait independensi observasi antara kelompok.
g. Ekstensi Model
   ANOVA memiliki ekstensi yang lebih beragam untuk analisis kompleks (seperti ANOVA faktorial, ANOVA pengukuran berulang) dibandingkan uji-t, memungkinkan analisis struktur data yang lebih rumit.

## 10.2 Konsep Dasar ANOVA (*Analisis Varians*)

### 10.2.1 Asumsi-asumsi dalam ANOVA

ANOVA, seperti banyak metode statistik parametrik lainnya, dibangun di atas beberapa asumsi fundamental yang harus dipenuhi untuk memastikan validitas hasil analisis (Cumming, 2014):

a. Independensi Observasi
   Setiap pengamatan harus independen dari pengamatan lainnya. Pelanggaran terhadap asumsi ini dapat terjadi dalam desain pengukuran berulang atau data berpasangan tanpa penyesuaian yang tepat.
b. Normalitas
   Distribusi residual (selisih antara nilai observasi dan nilai yang diprediksi oleh model) harus mengikuti distribusi normal. Asumsi ini dapat diperiksa melalui uji formal seperti *Shapiro-Wilk* atau metode grafik seperti Q-Q plot.
c. Homogenitas Varians (Homoskedastisitas)
   Varians di dalam setiap kelompok yang dibandingkan harus relatif sama. Asumsi ini dapat diuji dengan uji Levene atau uji Bartlett. Ketika asumsi ini dilanggar, modifikasi seperti ANOVA Welch atau transformasi data dapat dipertimbangkan.
d. Pengukuran pada Skala Interval atau Rasio
   Variabel dependen harus diukur pada skala interval atau rasio untuk menghasilkan interpretasi yang bermakna.
e. Tidak Adanya Outlier Signifikan
   Data tidak boleh mengandung outlier ekstrem yang dapat mendistorsi hasil analisis.

### 10.2.2 Varians Antar Kelompok dan Varians Dalam Kelompok

Konsep inti ANOVA adalah partisi varians total menjadi dua komponen (Pyzdek, 2021):

a. Varians Antar Kelompok (*Between-Group Variance*)
   Mengukur seberapa jauh rata-rata setiap kelompok berbeda dari rata-rata keseluruhan. Varians ini mencerminkan variabilitas yang disebabkan oleh perbedaan perlakuan atau faktor yang sedang diteliti. Semakin besar perbedaan antar kelompok, semakin besar varians antar kelompok.
b. Varians Dalam Kelompok (*Within-Group Variance*)
   Mengukur variabilitas pengamatan di dalam setiap kelompok atau disebut juga varians error. Varians ini mencerminkan variabilitas alami atau kebetulan yang tidak dapat dijelaskan oleh faktor yang sedang diteliti.

### 10.2.3 Hipotesis Nol dan Hipotesis Alternatif

Dalam ANOVA, pengujian hipotesis dirumuskan sebagai berikut (Huang et al., 2023):

a. Hipotesis Nol ($H_0$)
   Menyatakan bahwa tidak ada perbedaan yang signifikan antara rata-rata semua kelompok yang dibandingkan. Secara matematis, untuk k kelompok, $H_0$: $\mu_1 = \mu_2 = ... = \mu_k$, di mana $\mu_i$ adalah rata-rata populasi untuk kelompok ke-i.
b. Hipotesis Alternatif ($H_1$)
   Menyatakan bahwa setidaknya ada satu rata-rata kelompok yang berbeda secara signifikan dari yang lain. Secara matematis, $H_1$: tidak semua $\mu_i$ sama.

### 10.2.4 Nilai F dan Distribusi F

Statistik F adalah rasio antara *Mean Square Between* (MSB) dan *Mean Square Within* (MSW) (Thakkar, 2020):

$$F = MSB / MSW = (SSB/df_1) / (SSW/df_2)$$

di mana :

- ✓ SSB adalah *Sum of Squares Between* (jumlah kuadrat antar kelompok).
- ✓ SSW adalah *Sum of Squares Within* (jumlah kuadrat dalam kelompok).
- ✓ $df_1 = k - 1$ (derajat kebebasan untuk pembilang).
- ✓ $df_2 = N - k$ (derajat kebebasan untuk penyebut).
- ✓ k adalah jumlah kelompok.
- ✓ N adalah total jumlah pengamatan.

Statistik F mengikuti distribusi F dengan derajat kebebasan $df_1$ dan $df_2$. Nilai F yang besar mengindikasikan bahwa variabilitas antar kelompok lebih besar daripada variabilitas dalam kelompok, mendukung penolakan hipotesis nol.

Keputusan untuk menolak $H_0$ didasarkan pada perbandingan nilai F hitung dengan nilai F kritis dari tabel distribusi F pada tingkat signifikansi tertentu (α). Jika F hitung > F kritis, maka $H_0$ ditolak, menunjukkan bahwa setidaknya satu kelompok memiliki rata-rata yang berbeda secara signifikan.

Dalam era komputasi modern, p-value yang terkait dengan nilai F sering digunakan sebagai dasar keputusan. Jika p-value < α, maka $H_0$ ditolak.

## 10.3 Jenis-jenis ANOVA

### 10.3.1 ANOVA Satu Arah (*One-way* ANOVA)

ANOVA Satu Arah merupakan bentuk paling sederhana dari Analisis Varians yang digunakan ketika peneliti ingin membandingkan rata-rata dari tiga atau lebih kelompok independen berdasarkan satu faktor atau variabel independen kategoris (Hanafi Azman Ong et al., 2017). Misalnya, membandingkan efektivitas tiga jenis pupuk berbeda terhadap pertumbuhan tanaman.

Dalam *One-way ANOVA*, variasi total dalam data dipartisi menjadi dua komponen: variasi antar kelompok (between-group variation) yang menunjukkan efek dari perlakuan yang berbeda, dan variasi dalam kelompok (within-group variation) yang mencerminkan variabilitas alami atau error.

Model statistik untuk *One-way ANOVA* dapat dinyatakan sebagai:

$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$

di mana :

- ✓ $Y_{ij}$ adalah nilai observasi ke-i dalam kelompok j.
- ✓ $\mu$ adalah rata-rata keseluruhan populasi.
- ✓ $\alpha_j$ adalah efek dari perlakuan j.
- ✓ $\varepsilon_{ij}$ adalah error acak.

Hipotesis nol dalam *One-way ANOVA* menyatakan bahwa semua kelompok memiliki rata-rata populasi yang sama ($H_0$: $\mu_1 = \mu_2 = ... = \mu_k$), sementara hipotesis alternatif menyatakan bahwa setidaknya satu kelompok memiliki rata-rata yang berbeda.

### 10.3.2 ANOVA Dua Arah (*Two-way* ANOVA)

ANOVA Dua Arah memperluas analisis dengan mempertimbangkan dua faktor atau variabel

independen kategoris secara simultan (MUT et al., 2019). Pendekatan ini memungkinkan peneliti tidak hanya menguji efek utama dari setiap faktor tetapi juga interaksi antara kedua faktor tersebut.

Misalnya, seorang peneliti pertanian mungkin ingin menyelidiki pengaruh jenis pupuk (faktor A) dan frekuensi penyiraman (faktor B) terhadap hasil panen. *Two-way ANOVA* akan mengevaluasi :

a. Efek utama faktor A (perbedaan rata-rata hasil panen antar jenis pupuk).
b. Efek utama faktor B (perbedaan rata-rata hasil panen antar frekuensi penyiraman).
c. Efek interaksi A×B (apakah pengaruh jenis pupuk bervariasi berdasarkan frekuensi penyiraman).

Model statistik untuk *Two-way ANOVA* dapat dinyatakan sebagai:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

di mana :

- ✓ $Y_{ijk}$ adalah observasi ke-k dalam sel kombinasi faktor i dan j.
- ✓ $\mu$ adalah rata-rata keseluruhan.
- ✓ $\alpha_i$ adalah efek faktor A pada level i.
- ✓ $\beta_j$ adalah efek faktor B pada level j.
- ✓ $(\alpha\beta)_{ij}$ adalah efek interaksi antara faktor A level i dan faktor B level j.
- ✓ $\varepsilon_{ijk}$ adalah error acak.

*Two-way ANOVA* dapat dilakukan dengan atau tanpa replikasi, meskipun versi dengan replikasi (*multiple observations per cell*) diperlukan untuk menguji interaksi.

### 10.3.3 ANOVA Multifaktor

ANOVA Multifaktor merupakan perluasan dari *Two-way ANOVA* yang melibatkan tiga atau lebih variabel independen kategoris (faktor) (Tian & Zhou, 2023). Misalnya, seorang peneliti mungkin menyelidiki efek dari jenis pupuk, frekuensi penyiraman, dan intensitas cahaya terhadap hasil panen.

Dalam ANOVA Multifaktor, kita dapat menguji:

a. Efek utama dari setiap faktor.
b. Interaksi dua arah antara setiap pasangan faktor.
c. Interaksi tiga arah atau lebih antara kombinasi faktor.

Model statistik ANOVA Multifaktor dengan tiga faktor dapat dinyatakan sebagai:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijkl}$$

Kompleksitas model meningkat secara eksponensial dengan bertambahnya jumlah faktor, menghasilkan banyak efek interaksi yang mungkin sulit diinterpretasikan. Oleh karena itu, prinsip parsimoni sering diterapkan, di mana peneliti fokus pada interaksi tingkat rendah (biasanya dua atau tiga arah) yang memiliki makna praktis atau teoretis.

ANOVA Multifaktor membutuhkan ukuran sampel yang lebih besar untuk mempertahankan power statistik yang memadai, terutama untuk mendeteksi interaksi kompleks.

### 10.3.4 ANOVA dengan Pengukuran Berulang (*Repeated Measures* ANOVA)

ANOVA dengan Pengukuran Berulang digunakan ketika pengukuran diambil dari subjek yang sama pada beberapa kondisi berbeda atau titik waktu (Dimitrov &

Svetlana, 2014). Pendekatan ini meningkatkan efisiensi statistik dengan mengendalikan variabilitas antar subjek dan memungkinkan ukuran sampel yang lebih kecil.

Contoh klasik adalah pengukuran tekanan darah pasien sebelum, selama, dan setelah perawatan, atau mengukur performa subjek dalam beberapa kondisi eksperimental berbeda.

Berbeda dengan ANOVA standar yang mengasumsikan independensi pengamatan, Repeated Measures ANOVA mengatasi korelasi antar pengukuran dari subjek yang sama dengan mempartisi variabilitas dalam subjek (*within-subject variability*) dari variabilitas antar subjek (*between-subject variability*).

Model statistik untuk *One-way Repeated Measures ANOVA* dapat dinyatakan sebagai:

$$Y_{ij} = \mu + \pi_i + \tau_j + \varepsilon_{ij}$$

di mana :

- ✓ $Y_{ij}$ adalah pengamatan dari subjek i pada kondisi j.
- ✓ $\mu$ adalah rata-rata keseluruhan.
- ✓ $\pi_i$ adalah efek subjek i (variabilitas antar subjek).
- ✓ $\tau_j$ adalah efek kondisi j (variabel independen yang dimanipulasi).
- ✓ $\varepsilon_{ij}$ adalah error acak.

*Repeated Measures ANOVA* memiliki asumsi tambahan yang disebut sphericity—asumsi bahwa varians dari perbedaan antar semua pasangan kondisi adalah sama. Ketika asumsi ini dilanggar, koreksi seperti *Greenhouse-Geisser* atau *Huynh-Feldt* sering diterapkan.

ANOVA Pengukuran Berulang dapat diperluas untuk desain faktorial yang melibatkan faktor *within-subject* (subjek mengalami semua level) dan *between-subject* (subjek hanya mengalami satu level), menghasilkan desain *Split-Plot* atau *Mixed Design ANOVA*

yang menggabungkan aspek analisis berulang dan independen.

## 10.4 Langkah-langkah Perhitungan ANOVA

Perhitungan ANOVA merupakan proses sistematis yang melibatkan beberapa komponen utama untuk menguji signifikansi perbedaan antara rata-rata kelompok.

### 10.4.1 Perhitungan Jumlah Kuadrat (*Sum of Squares*)

Jumlah Kuadrat merupakan ukuran variabilitas yang mendasari ANOVA (Faturachman & Mustafa, 2012). Terdapat tiga komponen Jumlah Kuadrat yang perlu dihitung:

a. *Total Sum of Squares* (SST)
   Mengukur variasi total dalam data tanpa mempertimbangkan pengelompokan.
   $$SST = \Sigma(Y_{ij} - \bar{Y})^2$$
   di mana $Y_{ij}$ adalah pengamatan individual dan $\bar{Y}$ adalah rata-rata keseluruhan dari semua pengamatan.

b. *Between-group Sum of Squares* (SSB)
   Mengukur variasi yang disebabkan oleh perbedaan antara rata-rata kelompok.
   $$SSB = \Sigma n_j(\bar{Y}_j - \bar{Y})^2$$
   di mana $n_j$ adalah jumlah pengamatan dalam kelompok j, dan $\bar{Y}_j$ adalah rata-rata kelompok j.

c. *Within-group Sum of Squares* (SSW)
   Juga disebut *Error Sum of Squares* (SSE), mengukur variasi acak dalam setiap kelompok.
   $$SSW = \Sigma(Y_{ij} - \bar{Y}_j)^2$$
   atau dapat dihitung sebagai SSW = SST – SSB

Identitas penting dalam ANOVA adalah: SST = SSB + SSW.

Partisi ini memungkinkan kita untuk memisahkan variasi yang dapat dijelaskan oleh faktor (SSB) dari variasi acak (SSW).

### 10.4.2 Derajat Kebebasan (*Degrees of Freedom*)

Derajat kebebasan mengindikasikan jumlah nilai independen yang dapat bervariasi dalam perhitungan statistik (Purwanto et al., 2021):

d. *Total Degrees of Freedom* ($df_1$)
Jumlah total pengamatan dikurangi 1.
$$df_1 = N - 1$$
di mana N adalah jumlah total pengamatan.

e. *Between-group Degrees of Freedom* ($df_2$)
Jumlah kelompok dikurangi 1.
$$df_2 = k - 1$$
di mana k adalah jumlah kelompok.

f. *Within-group Degrees of Freedom* ($df_3$)
*Total degrees of freedom* dikurangi between-group degrees of freedom.
$$df_3 = df_1 - df_2 = N - k$$
Seperti halnya jumlah kuadrat, derajat kebebasan juga memenuhi relasi: $df_1 = df_2 + df_3$

### 10.4.3 Rata-rata Kuadrat (*Mean Square*)

Rata-rata kuadrat diperoleh dengan membagi jumlah kuadrat dengan derajat kebebasan yang sesuai (Purwanto A. et al., 2020):

g. *Mean Square Between* (MSB)
Estimasi varians antar kelompok.
$$MSB = SSB / df_2 = SSB / (k - 1)$$

h. *Mean Square Within* (MSW)
Estimasi varians dalam kelompok atau varians error.
$$MSW = SSW / df_3 = SSW / (N - k)$$
Jika hipotesis nol benar (tidak ada perbedaan antar kelompok), MSB dan MSW keduanya mengestimasi varians populasi yang sama. Namun, jika terdapat perbedaan nyata antar kelompok, MSB akan cenderung lebih besar daripada MSW.

### 10.4.4 Rasio F

Rasio F adalah statistik uji dalam ANOVA yang dihitung sebagai rasio dari MSB terhadap MSW (Huang et al., 2023):

$$F = MSB / MSW$$

Rasio F mengikuti distribusi F dengan derajat kebebasan $df_2$ dan $df_3$. Nilai F yang besar mengindikasikan bahwa varians antar kelompok secara signifikan lebih besar daripada varians dalam kelompok, memberikan bukti melawan hipotesis nol.

Untuk mengevaluasi signifikansi statistik dari rasio F, kita membandingkannya dengan nilai kritis dari tabel distribusi F pada tingkat signifikansi tertentu ($\alpha$) dan derajat kebebasan yang sesuai. Jika F hitung > F kritis, kita menolak hipotesis nol, menunjukkan bahwa setidaknya satu kelompok memiliki rata-rata yang berbeda secara signifikan.

Dalam penerapan modern, software statistik biasanya menghitung *p-value* yang terkait dengan nilai F. Jika *p-value* < $\alpha$, kita menolak hipotesis nol.

### 10.4.5 Tabel ANOVA

Hasil perhitungan ANOVA biasanya disajikan dalam format tabel standar yang merangkum semua komponen penting (MUT et al., 2019):

| ***Sumber Variasi*** | ***Sum of Squares (SS)*** | ***Degrees of Freedom (df)*** | ***Mean Square (MS)*** | ***F*** | ***p-value*** |
|---|---|---|---|---|---|
| Antar Kelompok | SSB | k - 1 | MSB = SSB/(k-1) | F = MSB/MSW | p |
| Dalam Kelompok | SSW | N - k | MSW = SSW/(N-k) | | |
| Total | SST | N - 1 | | | |

Tabel ANOVA memberikan ringkasan komprehensif dari hasil analisis, memudahkan interpretasi dan pelaporan. Elemen-elemen penting dalam tabel ini termasuk:

a. Sumber Variasi
   Mengidentifikasi komponen variasi yang dianalisis.
b. *Sum of Squares*
   Menunjukkan jumlah variabilitas yang terkait dengan setiap sumber.
c. *Degrees of Freedom*
   Mengindikasikan jumlah nilai bebas dalam perhitungan.
d. *Mean Square*
   Mengukur variasi rata-rata per derajat kebebasan.
e. *F-value*
   Statistik uji yang digunakan untuk mengevaluasi signifikansi perbedaan antar kelompok.
f. *p-value*
   Probabilitas memperoleh hasil yang diamati jika hipotesis nol benar.

Untuk ANOVA yang lebih kompleks, seperti *Two-way ANOVA*, tabel akan mencakup baris tambahan untuk efek utama dari faktor tambahan dan interaksi antar faktor.

Interpretasi tabel ANOVA berfokus terutama pada nilai F dan p-value terkait. Jika p-value kurang dari tingkat signifikansi yang dipilih (biasanya 0.05), kita dapat menyimpulkan bahwa terdapat perbedaan signifikan di antara kelompok-kelompok yang dibandingkan.

## 10.5 Uji Lanjut (*Post Hoc Tests*)

Uji lanjut atau post hoc tests merupakan serangkaian prosedur statistik yang dilakukan setelah ANOVA menghasilkan hasil yang signifikan (Hanafi Azman Ong et al., 2017). Tujuan utama dari uji lanjut adalah untuk menentukan secara spesifik kelompok mana yang berbeda secara signifikan dari yang lain. Hal ini diperlukan karena ANOVA hanya menginformasikan bahwa terdapat perbedaan signifikan di antara kelompok-kelompok, namun tidak mengidentifikasi di mana perbedaan tersebut terletak.

### 10.5.1 Uji Tukey HSD (*Honestly Significant Difference*)

Uji Tukey HSD dikembangkan oleh John Tukey dan merupakan salah satu uji post hoc yang paling popular (Dimitrov & Svetlana, 2014). Uji ini mengendalikan tingkat kesalahan familywise (*familywise error rate*) ketika melakukan multiple comparisons.

1. Prinsip Dasar
   Tukey HSD membandingkan semua pasangan kelompok yang mungkin dan menentukan nilai kritis

untuk perbedaan rata-rata berdasarkan distribusi studentized range. Perbedaan rata-rata yang melebihi nilai kritis dianggap signifikan secara statistik.

2. Formula

   $$HSD = q(\alpha, k, N-k) \times \sqrt{(MSW/n)}$$

   di mana :
   - ✓ $q(\alpha, k, N-k)$ adalah nilai kritis dari distribusi *studentized range* pada tingkat signifikansi $\alpha$.
   - ✓ k adalah jumlah kelompok.
   - ✓ N adalah jumlah total observasi.
   - ✓ MSW adalah *Mean Square Within* dari ANOVA.
   - ✓ n adalah jumlah observasi per kelompok (diasumsikan sama).
3. Keunggulan
   - ✓ Menyediakan keseimbangan yang baik antara menghindari kesalahan Tipe I dan mempertahankan power statistik.
   - ✓ Sangat direkomendasikan untuk perbandingan semua pasangan (*pairwise comparisons*).
   - ✓ Relatif mudah diinterpretasikan.
4. Keterbatasan
   - ✓ Mengasumsikan ukuran sampel yang sama untuk semua kelompok (meskipun versi modifikasi tersedia untuk ukuran sampel yang tidak sama).
   - ✓ Kurang tepat untuk pengujian kompleks seperti perbandingan subset kelompok.

### 10.5.2 Uji Bonferroni

Uji Bonferroni merupakan metode sederhana namun konservatif untuk mengendalikan tingkat

kesalahan Tipe I dalam *multiple comparisons* (Dabas, 2024). Uji ini menerapkan penyesuaian (*adjustment*) terhadap tingkat signifikansi berdasarkan jumlah perbandingan yang dilakukan.

a. Prinsip Dasar
   Bonferroni membagi tingkat signifikansi α dengan jumlah perbandingan untuk mendapatkan tingkat signifikansi yang disesuaikan (*adjusted significance level*).
b. Formula
   $$\alpha' = \alpha/m$$
   di mana :
   - ✓ α' adalah tingkat signifikansi yang disesuaikan.
   - ✓ α adalah tingkat signifikansi keseluruhan (biasanya 0.05).
   - ✓ m adalah jumlah perbandingan (untuk k kelompok, m = k(k-1)/2 untuk semua pasangan perbandingan)
   - ✓ Setiap perbandingan diuji menggunakan *t-test* dengan tingkat signifikansi α'.

c. Keunggulan
   - ✓ Sangat sederhana untuk diterapkan.
   - ✓ Sangat ketat dalam mengendalikan tingkat kesalahan Tipe I.
   - ✓ Dapat digunakan untuk berbagai jenis perbandingan, tidak terbatas pada *pairwise comparisons*.

d. Keterbatasan
   - ✓ Sangat konservatif, terutama ketika jumlah perbandingan besar.
   - ✓ Power statistik rendah dibandingkan dengan metode lain.

- ✓ Meningkatkan risiko kesalahan Tipe II (tidak mendeteksi perbedaan yang sebenarnya ada).

### 10.5.3 Uji Scheffé

Uji *Scheffé* dikembangkan oleh Henry Scheffé dan merupakan uji *post hoc* yang paling fleksibel dan konservatif. Uji ini tidak hanya membandingkan pasangan kelompok namun juga memungkinkan pengujian kontras yang lebih kompleks (Bulman & Osborn, 1989).

a. Prinsip Dasar
   Scheffé menggunakan distribusi F untuk menentukan nilai kritis untuk semua kemungkinan kontras linear di antara rata-rata kelompok.

b. Formula
   $$S = \sqrt{[(k-1)F(\alpha, k-1, N-k) \times MSW \times (1/n_1 + 1/n_2)]}$$
   di mana :
   - ✓ $F(\alpha, k-1, N-k)$ adalah nilai kritis dari distribusi F.
   - ✓ k adalah jumlah kelompok.
   - ✓ N adalah jumlah total observasi.
   - ✓ MSW adalah *Mean Square Within* dari ANOVA.
   - ✓ $n_1$ dan $n_2$ adalah ukuran sampel dari kelompok yang dibandingkan.

c. Keunggulan
   - ✓ Sangat fleksibel, memungkinkan pengujian kontras kompleks.
   - ✓ Memberikan kontrol ketat terhadap *familywise error rate* untuk semua jenis kontras.
   - ✓ Dapat menangani ukuran sampel yang tidak sama.

d. Keterbatasan
   - ✓ Sangat konservatif, menghasilkan power statistik yang rendah.
   - ✓ Sulit diinterpretasikan untuk non-statistisi.
   - ✓ Dapat menghasilkan hasil yang berbeda dari uji lain yang lebih spesifik untuk *pairwise comparisons*.

### 10.5.4 Uji *Least Significant Difference* (LSD)

Uji LSD, yang dikembangkan oleh Ronald Fisher, merupakan uji *post hoc* paling sederhana dan paling kurang konservatif (Gut, 2009). Uji ini pada dasarnya adalah serangkaian *t-tests* tanpa penyesuaian tingkat signifikansi.

a. Prinsip Dasar
   LSD menguji setiap pasangan kelompok menggunakan *t-test* dengan MSW dari ANOVA sebagai estimasi varians.

b. Formula
   $$LSD = t(\alpha/2, N-k) \times \sqrt{(MSW \times (1/n_1 + 1/n_2))}$$
   di mana :
   - ✓ $t(\alpha/2, N-k)$ adalah nilai kritis dari distribusi t.
   - ✓ N adalah jumlah total observasi.
   - ✓ k adalah jumlah kelompok.
   - ✓ MSW adalah *Mean Square Within* dari ANOVA.
   - ✓ $n_1$ dan $n_2$ adalah ukuran sampel dari kelompok yang dibandingkan.

c. Keunggulan
   - ✓ Memiliki power statistik yang tinggi
   - ✓ Sederhana untuk diterapkan dan diinterpretasikan.
   - ✓ Dapat menangani ukuran sampel yang tidak sama.

d. Keterbatasan
   - ✓ Tidak mengendalikan *familywise error rate*, sehingga risiko kesalahan Tipe I tinggi.
   - ✓ Umumnya hanya direkomendasikan ketika ANOVA menunjukkan hasil signifikan dan jumlah kelompok kecil.
   - ✓ Tidak cocok untuk perbandingan yang banyak.

## 10.6 Ukuran Efek dalam ANOVA

Ukuran efek (*effect size*) merupakan komponen penting dalam analisis statistik yang menginformasikan besarnya perbedaan atau kekuatan hubungan antar variabel (Nasution et al., 2024). Dalam konteks ANOVA, ukuran efek memungkinkan peneliti untuk mengevaluasi signifikansi praktis dari hasil yang ditemukan, melengkapi informasi signifikansi statistik yang diperoleh dari nilai p.

**1. *Eta squared* ($\eta^2$)**

*Eta squared* merupakan ukuran efek yang paling sederhana dan sering digunakan dalam ANOVA. *Eta squared* mengukur proporsi varians total yang dapat dijelaskan oleh faktor atau variabel independen (Pyzdek, 2021).

a. Formula

$$\eta^2 = SSbetween / Sstotal$$

di mana :
   - ✓ *SSbetween* adalah jumlah kuadrat antar kelompok.
   - ✓ *SStotal* adalah jumlah kuadrat total.

b. Interpretasi

*Eta squared* memiliki rentang nilai dari 0 hingga 1, dengan :
   - ✓ $\eta^2 \approx 0.01$ menunjukkan efek kecil.

- ✓ $\eta^2 \approx 0.06$ menunjukkan efek sedang.
- ✓ $\eta^2 \approx 0.14$ menunjukkan efek besar.
- ✓ Nilai 0.20 berarti bahwa 20% dari varians dalam variabel dependen dapat dijelaskan oleh perbedaan antar kelompok.

c. Keunggulan
- ✓ Mudah dihitung dan diinterpretasikan.
- ✓ Memberikan persentase varians yang dapat dijelaskan secara langsung.
- ✓ Banyak *software* statistik secara otomatis melaporkan nilai ini.

d. Keterbatasan
- ✓ Cenderung memberikan estimasi yang bias (overestimasi) terhadap ukuran efek populasi, terutama dengan ukuran sampel kecil.
- ✓ Tidak dapat dibandingkan secara langsung antar studi dengan desain yang berbeda.
- ✓ Tidak mempertimbangkan kompleksitas desain penelitian.

**2. *Omega squared* ($\omega^2$)**

*Omega squared* merupakan alternatif dari eta squared yang memberikan estimasi yang kurang bias terhadap ukuran efek dalam populasi. *Omega squared* mengoreksi bias dalam eta squared dengan mempertimbangkan derajat kebebasan dan *mean square error* (Thakkar, 2020).

a. Formula

$\omega^2$ = (*SSbetween* - (k-1) × *MSwithin*) / (*SStotal* + *MSwithin*)

di mana :
- ✓ *SSbetween* adalah jumlah kuadrat antar kelompok.
- ✓ *SStotal* adalah jumlah kuadrat total.

- ✓ k adalah jumlah kelompok.
- ✓ *MSwithin* adalah rata-rata kuadrat dalam kelompok

b. Interpretasi

*Omega squared* juga memiliki rentang nilai dari 0 hingga 1, dengan pedoman interpretasi yang serupa dengan *eta squared*:

- ✓ $\omega^2 \approx 0.01$ menunjukkan efek kecil.
- ✓ $\omega^2 \approx 0.06$ menunjukkan efek sedang.
- ✓ $\omega^2 \approx 0.14$ menunjukkan efek besar.
- ✓ Namun, nilai omega squared biasanya lebih kecil daripada eta squared untuk kumpulan data yang sama.

c. Keunggulan

- ✓ Memberikan estimasi yang kurang bias dibandingkan eta squared.
- ✓ Lebih stabil antar sampel.
- ✓ Lebih tepat untuk mengestimasi ukuran efek dalam populasi.

d. Keterbatasan

- ✓ Lebih kompleks untuk dihitung.
- ✓ Tidak semua software statistik menyediakan omega squared secara otomatis.
- ✓ Dapat menghasilkan nilai negatif dalam kondisi tertentu, yang sulit diinterpretasikan.

### *3. Cohen's f*

*Cohen's f* merupakan ukuran efek alternatif yang diperkenalkan oleh Jacob Cohen. Berbeda dengan eta squared dan omega squared yang mengukur proporsi varians yang dijelaskan, Cohen's f menyatakan efek dalam bentuk deviasi standar (Tian & Zhou, 2023).

a. Formula

$$f = \sqrt{(\eta^2 / (1 - \eta^2))}$$

Atau dihitung langsung dari statistik F

$$f = \sqrt{(F \times (k-1) / N)}$$

di mana :

- ✓ $\eta^2$ adalah eta squared.
- ✓ F adalah nilai F dari ANOVA.
- ✓ k adalah jumlah kelompok.
- ✓ N adalah jumlah total observasi.

b. Interpretasi

Cohen's f memiliki rentang nilai dari 0 hingga tak terhingga, dengan konvensi interpretasi:

- ✓ f = 0.10 menunjukkan efek kecil.
- ✓ f = 0.25 menunjukkan efek sedang.
- ✓ f = 0.40 menunjukkan efek besar.

c. Keunggulan

- ✓ Memungkinkan perbandingan langsung dengan ukuran efek lain yang menggunakan deviasi standar.
- ✓ Cocok untuk analisis power dan penentuan ukuran sampel.
- ✓ Tidak dibatasi oleh nilai maksimum 1.0, memberikan resolusi lebih baik untuk efek yang sangat besar.

d. Keterbatasan

- ✓ Kurang intuitif untuk diinterpretasikan dibandingkan ukuran efek berbasis varians.
- ✓ Tidak langsung mengindikasikan proporsi varians yang dijelaskan.
- ✓ Nilai dapat menjadi sangat besar untuk efek yang sangat kuat.

## 10.7 Penerapan ANOVA dengan *Software* Statistik

Analisis varians (ANOVA) dapat diimplementasikan dengan berbagai *software* statistik, dengan tiga pilihan populer: SPSS, R, dan Python. Masing-masing software memiliki keunggulan dan pendekatan yang berbeda dalam melakukan analisis ANOVA.

**1. ANOVA dengan SPSS**

SPSS (*Statistical Package for the Social Sciences*) merupakan software statistik komersial yang banyak digunakan dalam penelitian sosial dan pendidikan. SPSS menyediakan antarmuka pengguna grafis (GUI) yang intuitif untuk melakukan ANOVA (Purwanto et al., 2021).

a. Langkah-langkah dasar melakukan ANOVA di SPSS
   - ✓ Pilih menu *Analyze > Compare Means > One-Way ANOVA* (untuk ANOVA satu arah).
   - ✓ Pindahkan variabel dependen ke kotak "*Dependent List*" dan variabel independen ke kotak "*Factor*".
   - ✓ Klik "*Post Hoc*" untuk memilih uji lanjut (misalnya Tukey, Bonferroni).
   - ✓ Klik "*Options*" untuk menambahkan statistik deskriptif, uji homogenitas, dan ukuran efek.
   - ✓ Klik "OK" untuk menjalankan analisis.

SPSS secara otomatis menghasilkan tabel ANOVA lengkap dengan nilai F, *p-value*, dan ukuran efek (*eta squared*). Untuk ANOVA faktorial, peneliti dapat menggunakan menu *Analyze* > *General Linear Model* > *Univariate* yang memungkinkan analisis lebih kompleks termasuk interaksi antar faktor.

Kelebihan SPSS terletak pada kemudahan penggunaan dan interpretasi hasil, membuatnya ideal untuk peneliti tanpa latar belakang pemrograman yang kuat.

## DAFTAR PUSTAKA

Brown, E. N., & Kass, R. E. (2009). Special section: statistical training and curricular revision what is statistics? *American Statistician*, *63*(2), 105–110. https://doi.org/10.1198/tast.2009.0019

Bulman, J. S., & Osborn, J. F. (1989). Descriptive statistics. *British Dental Journal*, *166*(2), 51–54. https://doi.org/10.1038/sj.bdj.4806708

Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Dabas, P. (2024). *Descriptive and Inferential Statistics Descriptive and Inferential Statistics Using R Hands-On Examples with Practice Exercises and Solutions Sultan Chand & Sons Scan to Preview*.

Dimitrov, D., & Svetlana, V. (2014). Some Opportunities of CMS OpenCart to Realization of e-Shop. *4th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education*, *100*, 413–420.

Faturachman, D., & Mustafa, S. (2012). Performance of Safety Sea Transportation. *Procedia - Social and Behavioral Sciences*, *57*, 368–372. https://doi.org/10.1016/j.sbspro.2012.09.1199

Gut, A. (2009). *Texts in Statistics: Intermediate course in probability*.

Hanafi Azman Ong, M., Puteh, F., Teknologi MARA, U., & Alam Selangor, S. (2017). Quantitative Data Analysis: Choosing Between SPSS, PLS and AMOS in Social

Science Research. *International Interdisciplinary Journal of Scientific Research*, *3*(1), 1–12. www.iijsr.org

Huang, J. F., Chen, C. T. A., Chen, M. H., Huang, S. L., & Hsu, P. Y. (2023). Structural Equation Modeling of the Marine Ecological System in Nanwan Bay Using SPSS Amos. *Sustainability (Switzerland)*, *15*(14). https://doi.org/10.3390/su151411435

MUT, A. İ., KUTLUCA, T., & GEÇİCİ, M. E. (2019). Investigation of the Articles Published in the Journal of 'Education and Science' between 2011 & 2016 in the Context of the Use of SPSS and AMOS. *Eğitimde Kuram ve Uygulama*, *15*(1), 37–46. https://doi.org/10.17244/eku.465183

Nasution, A. S., Fajrin, N., Yusuf, M., Kurniawan, A., Bulan, D. D., Madubun, F. M., Satyahadewi, N., Nugraha, D., & Zoraida, M. N. (2024). *Statistika elementer*.

Purwanto, A., Asbari, M., Iman Santoso, T., Sunarsi, D., Ilham, D., Pamulang, U., Selatan, T., Palopo, I., & Selatan, S. (2021). Education Research Quantitative Analysis for Little Respondents: Comparing of Lisrel, Tetrad, GSCA, Amos, SmartPLS, WarpPLS, and SPSS. *Jurnal Studi Guru Dan Pembelajaran*, *4*(2), 335–350. https://e-journal.my.id/jsgp/article/view/1326

Purwanto A., Asbari, M., Santoso, T. I., Paramarta, V., & Sunarsi, D. (2020). Social and Management Research Quantitative analysis for Medium Sample. *Jurnal Ilmiah Ilmu Administrasi Publik.*, *9*(2), 518–532.

Pyzdek, T. (2021). Descriptive Statistics. *Management for Professionals*, *Part F458*, 145–149. https://doi.org/10.1007/978-3-030-69901-7_12

Thakkar, J. J. (2020). Structural equation modelling: Application for research and practice (with AMOS and R). In *Studies in Systems, Decision and Control* (Vol. 285). https://doi.org/10.1007/978-981-15-3793-6_1

Tian, L., & Zhou, X. (2023). An Empirical Study on the Causes of Mobile Phone Dependence of College Students. In *International Journal of Wireless Information Networks* (Vol. 30, Issue 1). Springer Singapore. https://doi.org/10.1007/s10776-021-00532-9

.

# BAB 11
# KORELASI DAN REGRESI LINEAR

**Oleh Tedi Hartoyo dan Unang Atmaja**

## 11.1 Pendahuluan

Korelasi adalah hubungan (*relationship* atau *association*) antara dua atau lebih variabel, sehingga muncul istilah *bivariable* atau *multivariable*. Pentingnya mempelajari hubungan antara dua buah atau lebih variabel adalah untuk mengetahui (menentukan, meramal atau menilai) nilai-nilai variabel yang satu, jika diketahui nilai-nilai variabel yang lainnya. Dengan kata lain, nilai variabel yang satu diduga secara tidak langsung dari nilai variabel yang lainnya. Variabel yang nilainya diduga secara tidak langsung itu disebut variabel tidak bebas (*dependent variable*), sedangkan variabel yang lainnya, yang digunakan untuk pendugaan disebut variabel bebas (*independent variable*).

Jika variabel tidak bebas dinyatakan dengan Y dan variabel bebas dinyatakan dengan X, maka hubungan fungsional antara variabel tidak bebas dengan variabel bebas itu dinamakan regresi. Jadi regresi antara variabel tidak bebas (Y) dengan variabel bebas (X) adalah Y = f(X).

Beberapa contoh penerapannya antara lain:

1. Menduga hubungan antara konsumsi dengan pendapatan.
2. Mencari hubungan antara jumlah ekspor salak dengan PDRB (Produk Domestik Regional Bruto).
3. Menduga hubungan antara jumlah penduduk dengan permintaan suatu produk.

## 11 .2 Fungsi dan Regresi

Untuk mencari hubungan antara variabel tidak bebas dengan variabel bebas, harus dilakukan pengamatan terhadap sejumlah sampel atau populasi. Pengamatan tersebut merupakan pasangan nilai dari kedua variabel tersebut. Tampilan pasangan variabel dapat dilihat pada tabel di bawah ini:

| **Nomor Responden** | **Pendapatan (X)** | **Konsumsi (Y)** |
|---|---|---|
| 1 | $X_1$ | $Y_1$ |
| 2 | $X_2$ | $Y_2$ |
| 3 | $X_3$ | $Y_3$ |
| 4 | $X_4$ | $Y_4$ |
| . | . | . |
| . | . | . |
| . | . | . |
| N | $X_n$ | $Y_n$ |

Jika digambarkan dalam salib sumbu X dan Y, maka akan diperoleh gambaran yang dinamakan diagram titik (*scatter diagram* atau *scatter plot*). Seandainya antara kedua variabel tersebut ada hubungan yang erat, maka diagram titik yang terbentuk mempunyai arah atau tendensi yang jelas. Selanjutnya dengan menarik garis perataan dari titik-titik tadi, maka dapat dilukiskan hubungan antara variabel tidak bebas (Y) dengan variabel bebas (X), sehingga dapat ditaksir nilai Y = Yi tertentu melalui nilai X = Xi tertentu dengan menggunakan garis tersebut. Garis itu dinamakan garis regresi.

Dalam fungsi statistik dimana dinyatakan hubungan antara X dan Y, pada umumnya titik-titik pada diagram tidak tepat berada pada garis regresinya, berbeda dengan fungsi matematika. Jadi setiap titik mempunyai simpangan atau deviasi (*deviation*) terhadap garis regresinya. Besar simpangan pada setiap titik itu menunjukkan keeratan hubungan antara kedua variabel tersebut.

Perbedaan antara hubungan matematik (fungsi matematik) dan hubungan statistik (fungsi statistik) adalah fungsi matematik tidak mempunyai deviasi, sedangkan fungsi statistik mempunyai deviasi. Dalam fungsi statistik, tujuannya adalah mencari deviasi sekecil mungkin atau mencari jumlah kuadrat deviasi sekecil mungkin.

## 11.3 Jenis-Jenis Regresi

Berdasarkan jumlah variabel bebasnya, regresi dapat dibedakan menjadi:

1. Regresi sederhana (*simple regression*), terdiri dari satu variabel X dan satu variabel Y, dengan formulasi sebagai berikut: $Y = f(X)$ $Y = \beta_0 + \beta_1 X + e$
2. Regresi berganda (*Multiple regression*), terdiri dari dua atau lebih variabel X dan satu variabel Y, dengan formulasi sebagai berikut: $Y = f(X_1, X_2, \ldots X_n)$ $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + e$

Keterangan:

$\beta_0$ = perpotongan dengan sumbu Y (intercept)

$\beta$ atau $\beta_1$, $\beta_2$, ... = koefisien regresi

e = error term (galat)

Berdasarkan bentuk kurvanya, regresi dapat dibedakan menjadi:

1. Regresi linear
2. Regresi non linear
   a. Regresi kuadratik
   b. Regresi kubik
   c. Regresi eksponensial
   d. Regresi logaritma

## 11.4 Menentukan Bentuk Hubungan Fungsional

Pada garis besarnya ada dua cara untuk mencari atau menentukan bentuk hubungan fungsional antara variabel tidak bebas (Y) dengan variabel bebas (X) dari suatu populasi, yaitu:

1. Cara matematik
2. Cara grafik

Pada cara matematik, hubungan fungsional itu ditentukan dengan menggunakan hubungan ilmu matematik. Pada cara grafik, hubungan ditentukan secara visual dengan melihat tendensi diagram titiknya, dimana garis regresi ditarik dari diagram titik sesuai dengan arahnya dan biasa disebut dengan *free hand method*. Kedua cara tersebut mempunyai kelebihan dan kelemahan masing-masing.

**Kelebihan cara matematik** yaitu:

- Dapat merumuskan hubungan fungsional dengan pasti, sehingga mudah diinterpretasikan
- Hasil rumusan yang diperoleh sangat objektif

**Kelemahan cara matematik** yaitu:

- Perumusan secara matematik tidak selalu mudah untuk disesuaikan dengan data yang tersedia
- Metode ini terlalu kaku, sehingga untuk memperoleh suatu hasil yang memuaskan tidak selalu tercapai

Kelemahan tersebut dapat dihindari dengan penggunaan cara grafik, dimana hubungan fungsional tersebut digambar sesuai dengan sifat data aslinya. Namun, cara grafik memiliki kelemahan yaitu:

- Sangat subjektif sifatnya, karena setiap orang akan menarik garis regresi yang berbeda-beda

- Hanya berlaku untuk hubungan dua buah variabel saja
- Hubungan antara tiga atau lebih variabel tidak dapat digambarkan, karena membutuhkan tiga dimensi dan sukar untuk menggambar grafiknya
- Membutuhkan pengalaman yang banyak

### 11.4.1 Persamaan Regresi Linear

Regresi linear digunakan untuk memodelkan hubungan antara variabel tidak bebas (Y) dengan variabel bebas (X) dan dapat digunakan untuk menduga atau meramal, juga dapat melihat perkembangan variabel tidak bebas akibat variabel bebas.

1. Bentuk Hubungan:

   Jika data yang digunakan dalam bentuk populasi, maka bentuk hubungan adalah sebagai berikut:

   $$Y = \beta_0 + \beta_1 X + e$$

   dan dapat ditulis menjadi Y* (Y dugaan) sebagai berikut:

   $$\hat{Y} = \beta_0 + \beta_1 X$$

   Jika data yang digunakan dalam bentuk sampel, maka bentuk hubungan tersebut dapat ditulis sebagai berikut:

   $$Y = a + bX + e$$

   $$\hat{Y} = a + bX$$

   Dimana:

   Y = Variabel tidak bebas

   X = Variabel bebas

   $\beta_0$ = a = Konstanta

   $\beta_1$ = b = Koefisien regresi

e = Error term (galat)

2. Asumsi Regresi Linear

   Ada beberapa asumsi yang harus dipenuhi untuk hubungan dua variabel yaitu variabel X dan variabel Y, sebagai berikut:

   1. Pengukuran variabel X dianggap tanpa error, jadi X merupakan suatu invarian
   2. Data mengikuti distribusi normal. Untuk setiap variabel X tertentu, nilai Y yang bersangkutan dianggap mempunyai distribusi normal
   3. Homogenitas Ragam. Ragam berlaku untuk setiap nilai X dan disebut "*common variance*"

3. Menentukan Koefisien Regresi

   Untuk menentukan nilai penduga koefisien regresi ($\beta_1$) yaitu b dan koefisien arah atau intercept ($\beta_0$) yaitu a, maka digunakan formulasi sebagai berikut:

   $b = JHK_{(x\gamma)} / JK_{(x)}$

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

   dimana:

   $\bar{Y}$ = rata-rata Y

   $\bar{X}$ = rata-rata X

   b = koefisien regresi

   $JHK_{(x\gamma)}$ = Jumlah Hasil Kali X dan Y

   $JK_{(x)}$ = Jumlah Kuadrat X

4. Pengujian Hipotesis
   - Hipotesis:

     $H_0 : \beta = 0$

     $H_1 : \beta \neq 0$

   - Uji statistik:

     1) **Dengan Uji-t**

     $t = (b - \beta) / S_{(\beta)} = b / S_{(\beta)}$ $S_{(\beta)} = \sqrt{[(JK_{(Y)} - b \times JHK_{(xy)}) / ((n-2) \times JK_{(x)})]}$

     Keputusan:

     Jika $t_{(hitun)} < t_{(ta\beta el)}$, maka $H_0$ diterima

     Jika $t_{(hitun)} \geq t_{(ta\beta el)}$, maka $H_0$ ditolak

     2) Dengan Uji-F (Analisis Variansi - ANOVA)

| **Sumber Keragaman** | **db** | **JK** | **KT** | **F hitung** | **F tabel** |
|---|---|---|---|---|---|
| Regresi | 1 | $JK_{(re)}$ | $KT_{(re)}$ | $KT_{(re)}/KT_{(e)}$ | |
| Galat | N-2 | $JK_{(e)}$ | $KT_{(e)}$ | | |
| Total | N-1 | $JK_{(total)}$ | | | |

   - ✓ Jumlah Kuadrat:

     $JK_{(re)}$ (JK regresi) = $b \times JHK_{(xy)}$

     $JK_{(total)}$ (JK total) = $JK_{(Y)}$

     $JK_{(e)}$ (JK galat) = $JK_{(total)} - JK_{(re)}$

   - ✓ Kuadrat Tengah:

     $KT_{(re)}$ (KT regresi) = $JK_{(re)} / db_{(re)}$

     $KT_{(e)}$ (KT galat) = $JK_{(e)} / db_{(e)}$

   - ✓ F hitung = $KT_{(re)} / KT_{(e)}$ F tabel = $F_{(}\alpha, db=N-2_{)}$

- ✓ Keputusan:

  Jika $F_{(hitun)} < F_{(tabel)}$, maka $H_0$ diterima

  Jika $F_{(hitun)} \geq F_{(tabel)}$, maka $H_0$ ditolak

**Contoh Kasus: Ekspor Salak dan PDRB**

Terdapat seorang petani salak yang melakukan ekspor dari hasil produksinya ke daerah lain. Jumlah ekspor salak dipengaruhi oleh PDRB (Produk Domestik Regional Bruto) daerah ekspor yang dituju. Hubungan antara PDRB dengan jumlah ekspor salak ditunjukkan oleh fungsi sebagai berikut:

$$\mathbf{Y = f(X) = \beta_0 + \beta_1 X}$$

dimana:

Y = Ekspor Salak (dalam kuintal)

X = PDRB daerah tujuan ekspor

Data hipotetis tentang ekspor salak dan PDRB ditunjukkan dalam tabel berikut:

| Tahun | Y (Ekspor Salak) | X (PDRB) |
|---|---|---|
| 1997 | 58,5 | 892,4 |
| 1998 | 64,0 | 1063,4 |
| 1999 | 75,9 | 1171,1 |
| 2000 | 94,4 | 1306,6 |
| 2001 | 131,9 | 1412,9 |
| 2002 | 126,9 | 1528,8 |
| 2003 | 155,4 | 1702,2 |
| 2004 | 185,8 | 1899,5 |
| 2005 | 217,8 | 2127,6 |
| 2006 | 260,9 | 2368,5 |

Pertanyaan:

a. Buatkan persamaan regresinya?
b. Apakah persamaan tersebut di poin (a) dapat digunakan untuk menduga?
c. Hitung nilai koefisien korelasi dan determinasinya dan apa artinya?

**Proses Pengujian:**

**Hipotesis:**

$H_0$: β = 0 (Tidak ada hubungan linier antara PDRB dengan jumlah ekspor salak)

$H_1$: β ≠ 0 (Terdapat hubungan linier antara PDRB dengan jumlah ekspor salak)

Menghitung koefisien regresi dan titik potong (intercept):

| X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|
| 892,4 | 58,5 | 796.377,76 | 3.422,25 | 52.205,4 |
| 1063,4 | 64,0 | 1.130.819,56 | 4.096,00 | 68.057,6 |
| 1171,1 | 75,9 | 1.371.475,21 | 5.760,81 | 88.886,49 |
| 1306,6 | 94,4 | 1.707.203,56 | 8.911,36 | 123.343,04 |
| 1412,9 | 131,9 | 1.996.286,41 | 17.397,61 | 186.361,51 |
| 1528,8 | 126,9 | 2.337.229,44 | 16.103,61 | 194.004,72 |
| 1702,2 | 155,4 | 2.897.484,84 | 24.149,16 | 264.521,88 |
| 1899,5 | 185,8 | 3.608.100,25 | 34.521,64 | 352.927,10 |
| 2127,6 | 217,8 | 4.526.681,76 | 47.436,84 | 463.391,28 |
| 2368,5 | 260,9 | 5.609.792,25 | 68.068,81 | 617.941,65 |
| Jumlah | 1371,5 | 25.981.451,04 | 229.868,09 | 2.411.640,67 |
| Rata-rata | 137,15 | 2.598.145,10 | 22.986,81 | 241.164,07 |

**Menghitung koefisien regresi:**

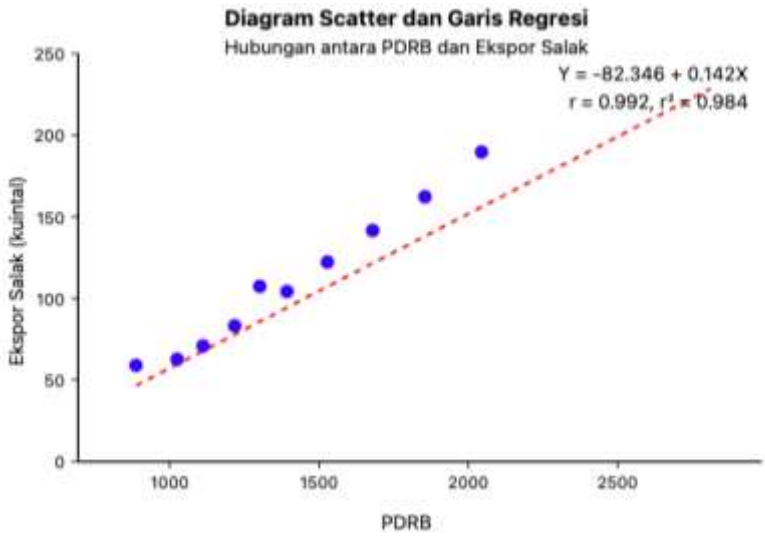
$b = JHK_{(xy)} / JK_{(x)} = [(\sum XY) - (\sum X)(\sum Y)/n] / [(\sum X^2) - (\sum X)^2/n]$
$= [2.411.640,67 - (15.473)(1.371,5)/10] / [25.981.451,04 - (15.473)^2/10] = 0,142$

Menghitung konstanta:

$a = \bar{Y} - b\bar{X} = 137,15 - (0,142)(1.547,3) = -82,346$

Jadi, persamaan regresinya adalah: $\hat{Y} = -82,346 + 0,142X$

Jika divisualisasikan, persamaan regresi ini akan menghasilkan garis seperti yang ditunjukkan pada grafik berikut:




Grafik di atas menunjukkan diagram titik (*scatter plot*) untuk data PDRB dan ekspor salak, serta garis regresi yang sesuai dengan persamaan di atas. Dapat dilihat bahwa titik-titik data cenderung tersebar dekat dengan garis regresi,

yang mengindikasikan hubungan yang kuat antara kedua variabel.

**Analisis Variansi (ANOVA):**

| Sumber Keragaman | db | JK | KT | F hitung | F tabel |
|---|---|---|---|---|---|
| Regresi | 1 | 41.087,195 | 41.087,195 | 483,614 | 7,57** |
| Galat | 8 | 679,670 | 84,959 | | |
| Total | 9 | 41.766,865 | | | |

**)sangat signifikan

Jumlah Kuadrat:

$JK_{(re)} = b \times JHK_{(xy)} = 0{,}142 \times [(\sum XY) - (\sum X)(\sum Y)/n] = 0{,}142 \times [2.411.640{,}67 - (15.473)(1.371{,}5)/10] = 41.087{,}195$

$JK_{(total)} = (\sum Y^2) - (\sum Y)^2/n = 229.868{,}09 - (1.371{,}5)^2/10 = 41.766{,}865$

$JK_{(e)} = JK_{(total)} - JK_{(re)} = 41.766{,}865 - 41.087{,}195 = 679{,}670$

Kuadrat Tengah:

$KT_{(re)} = JK_{(re)} / db_{(re)} = 41.087{,}195 / 1 = 41.087{,}195$

$KT_{(e)} = JK_{(e)} / db_{(e)} = 679{,}670 / 8 = 84{,}959$

F hitung $= KT_{(re)} / KT_{(e)} = 41.087{,}195 / 84{,}959 = 483{,}614$

**Uji-t:**

$S_{(\beta)} = \sqrt{(KT_{(e)} / JK_{(x)})} = \sqrt{[84{,}959 / ((\sum X^2) - (\sum X)^2/n)]} = \sqrt{[84{,}959 / (25.981.451{,}04 - (15.473)^2/10)]} = 0{,}00645$

t hitung $= b / S_{(\beta)} = 0{,}142 / 0{,}00645 = 21{,}991$**

**) t tabel pada α = 0,01 dan db = 8 sebesar 3,355 (sangat signifikan)

Keputusan: Karena $t_{(hitun)}$ (21,991) > $t_{(taβel)}$ (3,355), maka $H_0$ ditolak.

**Korelasi**

Besar kecilnya kekuatan hubungan antara variabel tidak bebas (Y) dengan variabel bebas (X) dinyatakan dengan suatu bilangan yang disebut koefisien korelasi (correlation coefficient). Koefisien korelasi untuk regresi sederhana biasa diberi simbol dengan huruf r dan ini digunakan untuk mengukur kekuatan atau keeratan hubungan antara variabel tidak bebas (Y) dengan variabel bebas (X). Kuadrat dari nilai koefisien korelasi yaitu biasa disebut koefisien determinasi ($r^2$). Nilai ini biasa diartikan sebagai ketepatan hubungan antara kedua variabel tersebut.

**Uji Korelasi**

Digunakan untuk mengukur kekuatan atau keeratan hubungan linear antara dua variabel dan tidak bisa digunakan untuk menduga. Bernilai antara -1 dan +1.

Hipotesis:

$H_0$: ρ = 0 (tidak ada korelasi antara kedua buah variabel)

$H_1$: ρ ≠ 0 (ada korelasi antara kedua buah variabel)

Cara menguji:

Menghitung nilai korelasi (r):

$r = JHK_{(xy)} / \sqrt{(JK_{(x)} \times JK_{(y)})}$

Uji statistik:

t hitung = $r \times \sqrt{[(n-2)/(1-r^2)]}$

Keputusan:

Jika $t_{(hitun)} < t_{(taβel)}$, maka $H_0$ diterima

Jika $t_{(hitun)} \geq t_{(taβel)}$, maka $H_0$ ditolak

**Contoh Aplikasi (Kasus Ekspor Salak):**

Hipotesis:

$H_0$: $\rho = 0$ (tidak ada korelasi antara kedua buah variabel)

$H_1$: $\rho \neq 0$ (ada korelasi antara kedua buah variabel)

Menghitung nilai korelasi (r):

$r = JHK_{(xy)} / \sqrt{(JK_{(x)} \times JK_{(y)})} = [(\sum XY) - (\sum X)(\sum Y)/n] / \sqrt{[((\sum X^2) - (\sum X)^2/n) \times ((\sum Y^2) - (\sum Y)^2/n)]} = [2.411.640,67 - (15.473)(1.371,5)/10] / \sqrt{[(25.981.451,04 - (15.473)^2/10) \times (229.868,09 - (1.371,5)^2/10)]} = 0,992$

Koefisien determinasi $(r^2) = (0,992)^2 = 0,984$ atau 98,4%

**Uji statistik:**

t hitung = $r \times \sqrt{[(n-2)/(1-r^2)]} = 0,992 \times \sqrt{[(10-2)/(1-(0,992)^2)]} = 0,992 \times \sqrt{[8/0,016]} = 0,992 \times \sqrt{500} = 0,992 \times 22,36 = 22,18$

Keputusan: Karena $t_{(hitun)}$ (22,18) > $t_{(taβel)}$ (3,355), maka $H_0$ ditolak.

**Kesimpulan:** Terdapat hubungan yang positif antara PDRB dengan jumlah ekspor salak, dengan keeratan hubungan sebesar 0,992 (99,2%), dan dilihat dari nilai koefisien

determinasinya yaitu sebesar 0,984 (98,4%), yang artinya bahwa 98,4% variasi ekspor salak dapat dijelaskan oleh variasi Produk Domestik Regional Bruto (PDRB).

## 11.5 Penutup

Analisis korelasi dan regresi linear merupakan metode statistik yang sangat berguna dalam bidang ekonomi dan bisnis untuk memahami dan memprediksi hubungan antar variabel. Bab ini telah membahas dasar-dasar korelasi dan regresi linear, mulai dari konsep dasar hingga aplikasi dalam kasus nyata.

Beberapa poin penting yang perlu diingat dari pembahasan ini antara lain:

1. Pembedaan Korelasi dan Regresi - Korelasi mengukur kekuatan dan arah hubungan antara variabel, sementara regresi memungkinkan kita untuk membangun model prediktif untuk memperkirakan nilai variabel tidak bebas berdasarkan variabel bebas.
2. Fungsi Korelasi - Koefisien korelasi (r) mengukur keeratan hubungan antara dua variabel, dengan nilai berkisar antara -1 hingga +1. Nilai mendekati +1 atau -1 menunjukkan hubungan yang kuat, sementara nilai mendekati 0 menunjukkan hubungan yang lemah.
3. Interpretasi Koefisien Determinasi - Koefisien determinasi ($r^2$) menunjukkan persentase variasi dalam variabel tidak bebas yang dapat dijelaskan oleh variabel bebas. Misalnya, $r^2$ = 0,984 berarti 98,4% variasi dalam variabel Y dapat dijelaskan oleh variabel X.
4. Pengujian Signifikansi - Uji t dan uji F digunakan untuk menentukan apakah hubungan yang diamati antara

variabel adalah signifikan secara statistik atau hanya terjadi karena kebetulan.

5. Asumsi Model Regresi - Validitas model regresi linear bergantung pada beberapa asumsi seperti linearitas, normalitas residual, dan homoskedastisitas (kesamaan varians).
6. Aplikasi Praktis - Regresi linear memiliki banyak aplikasi praktis dalam ekonomi dan bisnis, seperti menganalisis hubungan antara harga dan permintaan, memprediksi penjualan berdasarkan biaya iklan, atau mengestimasi pendapatan berdasarkan tingkat pendidikan.

Kemampuan untuk menganalisis hubungan antara variabel-variabel ekonomi dan bisnis melalui korelasi dan regresi linear merupakan keterampilan yang sangat berharga dalam pengambilan keputusan berdasarkan data. Namun, perlu diingat bahwa hubungan korelasional tidak selalu berarti hubungan kausal. Faktor-faktor lain yang tidak dimasukkan dalam model mungkin juga memengaruhi variabel tidak bebas.

Dalam praktiknya, analisis korelasi dan regresi seringkali merupakan langkah awal dalam analisis data yang lebih kompleks. Pemahaman yang baik tentang konsep-konsep dasar yang dibahas dalam bab ini akan membantu dalam mempelajari metode statistik lanjutan seperti regresi non-linear, analisis time series, atau analisis multivariat.

Ada beberapa saran yang dapat di ungkapkan dalam melakukan pengujian, bahwa jika pengujian hubungan (uji F atau uji-t) sudah dilakukan dan hasil nya signifikan, maka proses pencarian nilai koefisien korelasi atau koefisien determinasi dapat dilakukan, tetapi jika sebaliknya, bahwa

hasil pengujian hubungan tidak signifikan, maka tidak perlu dilakukan pencarian nilai koefisien korelasi dan determinasi.

## DAFTAR PUSTAKA

Snedecor G.W and Cochran W,G. 1971. *Statistical Method. The Iowa State University .*

*Press, Ames, Iowa, USA. Sixth Edition, Fourth Printing*.

Sudradjat S.W. M. 1984. Mengenal Ekonometrika Pemula. Penerbit Armico, Bandung.

Steel R.G.D and Torrie J.H. 1981. *Principles and Procedures of Statistics. A Biometrical*

*Approach. Mc.Graw-Hill International Book Company. International Student Edition. Second Edition*.

Walpole R.E. 1975. *Introduction to Statistics. New York Macmillan London Collier Macmillan*.

Walpole R.E. 1993. Pengantar Statistika. Gramedia Pustaka Utama. Indonesia.

# BAB 12
# REGRESI LINEAR BERGANDA

**Oleh Dodi**

## 12.1 Pendahuluan

Ketika terdapat variabel terikat dan variabel bebas, maka akan terjadi saling berpengaruh antara satu variabel dengan variabel lainnya. Untuk menganalisis bagaimana hubungannya, maka dikenal dengan istilah Regresi Linear. Secara pengertian, Regresi Linear adalah suatu metode statistik yang digunakan untuk menganalisis hubungan antara satu atau lebih variabel bebas dengan satu variabel terikat. Sedangkan Regresi Linear Berganda adalah suatu metode statistik yang digunakan untuk menganalisis hubungan antara satu variabel terikat dengan dua atau lebih variabel bebas.

Regresi linear berganda memiliki peran yang sangat berpengaruh dalam dunia ekonomi. Peran tersebut dirasakan ketika saat menganalisis data dan mengambil keputusan yang didasarkan pada fakta. Misalkan peran regresi linear berganda dalam peramalan penjualan dan permintaan. Kita dapat menganalisis faktor-faktor yang mempengaruhi penjualan, seperti harga, promosi, dan tren pasar serta membantu perusahaan dalam memperkirakan permintaan produk berdasarkan variabel ekonomi dan sosial. Efek tersebut tidak lepas dengan pengaruh harga terhadap

volume penjualan, yakni kita dapat mengidentifikasi faktor eksternal apa saja yang memengaruhi strategi harga. Dengan demikan, kita dapat mengidentifikasi faktor-faktor yang mempengaruhi keputusan pembelian pelanggan, seperti usia, pendapatan, dan preferensi produk serta membantu dalam menyusun strategi pemasaran yang lebih efektif berdasarkan data demografi dan psikografis pelanggan.

## 12.2 Konsep Dasar Linear Berganda

Regresi linear berganda dapat dinyatakan pada persamaan sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon$$

Keterangan:

$Y =$ variabel terikat

$X_1, X_2, X_n =$ variabel bebas

$\beta_0 =$ konstanta

$\beta_1, \beta_2, \beta_n =$ koefisien regresi yang menunjukan pengaruh masing-masing variabel bebas terhadap variabel terikat

$\varepsilon =$kesalahan atau gangguan dalam model

a. Variabel Terikat

Variabel terikat atau *dependent variable* adalah variabel yang dipengaruhi atau dijelaskan oleh satu atau lebih variabel bebas dalam suatu analisis statistik. Secara regresi, dapat dikatakan bahwa variabel terikat merupakan variabel yang menjadi fokus penelitian karena variabel tersebut diakibatkan dari pengaruh dari beberapa faktor.

Nilai variabel terikat dapat berubah yang diakibatkan dari pengaruh variabel bebas. Misalkan terdapat 1

variabel bebas saja yang mengalami perubahan nilai, maka nilai variabel terikat juga berubah. Nilai variabel terikat dapat berupa kuantitatif atau dapat berupa kualitatif. Jika nilainya bersifat kuantitatif berupa angka, misalnya nilai tes. Jika nilainya bersifat kualitatif, misalnya kepuasan responden.

Beberapa contoh variabel terikat ($Y$) adalah Penjualan Produk ($Y$) yang dipengaruhi oleh harga, dan diskon. Keuntungan suatu perusahaan ($Y$) dipengaruhi oleh jumlah pelanggan dan jumlah karyawan. Tingkat pengangguran ($Y$) dipengaruhi oleh inflasi dan pertumbuhan ekonomi.

b. Variabel Bebas

Variabel bebas atau *independent variable* adalah faktor yang mempengaruhi untuk menjelaskan variabel terikat. Dalam bidang ekonomi dan bisnis, variabel bebas digunakan untuk menganalisis dan memprediksi berbagai macam keadaan seperti kondisi bagaimana hasil penjualan suatu produk, bagaiaman keuntungan suatu perusahaan, serta inflasi dan pertumbuhan ekonomi pada suatu negara.

Variabel bebas (disimbolkan dengan $X$) sering digunakan untuk menganalisis faktor-faktor yang mempengaruhi indikator ekonomi utama seperti Produk Domestik Bruto (PDB), tingkat inflasi, dan tingkat pengangguran. Misalnya, kita ingin menganalisis faktor-faktor yang mempengaruhi PDB suatu negara. Beberapa variabel bebas yang dapat digunakan dalam regresi linear berganda adalah investasi ($X_1$), semakin tinggi investasi maka semakin tinggi pertumbuhan ekonomi. Variabel bebas lain yang mempengaruhi PDB adalah konsumsi masyarakat ($X_2$), konsumsi yang tinggi mendorong

peningkatan produksi. Variabel bebas lainnya yang mempengaruhi PDB adalah ekspor ($X_3$) dimana negara dengan ekspor tinggi cenderung memiliki PDB yang lebih besar daripada negara yang nilai ekspornya rendah.

c. Konstanta

Konstanta adalah nilai dari variabel terikat ($Y$) ketika semua variabel bebas ($X_1$, $X_2$, $X_n$) bernilai nol. Konstanta menunjukkan nilai awal dari variabel terikat saat semua faktor bebas tidak berpengaruh. Misalnya, dalam model pengiriman produk, jika $\beta_0$= 75, itu berarti saat tidak ada variabel bebas seperti promosi, harga, atau faktor lain, pengiriman dasar berjumlah 75 unit.

Namun, pada beberapa kasus, nilai $\beta_0$ bisa jadi tidak memiliki makna sebagaimana mestinya. Misalkan dalam model matematis yang memprediksi pendapatan berdasarkan pengalaman kerja dan pendidikan, nilai $\beta_0$ dapat bernilai negatif, yang dalam realitanya bahwa pendapatan tidak mungkin negatif.

d. Koefisien Regresi

Koefisien regresi dalam regresi linear berganda adalah angka yang menunjukkan seberapa besar pengaruh masing-masing variabel bebas terhadap variabel terikat. Koefisien ini menunjukkan arah antara variabel bebas dan variabel terikat. Koefisien regresi positif ($\beta$>0) apabila variabel bebas memiliki hubungan positif dengan variabel terikat (jika $X$ naik, maka $Y$ naik). Koefisien negatif ($\beta$<0) apabila variabel bebas memiliki hubungan negatif dengan variabel terikat (jika $X$ naik, maka $Y$ turun). Koefisien regresi juga menunjukkan besarnya hubungan antara variabel bebas dan variabel

terikat. Semakin besar nilai $\beta$ maka semakin kuat pengaruh variabel bebas terhadap variabel terikat.

Misalkan, faktor-faktor yang mempengaruhi PDB suatu negara adalah investasi, konsumsi, dan ekspor. Maka, kita ingin mengetahui bagaimana investasi ($X_1$), konsumsi ($X_2$), dan ekspor ($X_3$) mempengaruhi PDB ($Y$).

Model regresinya:

$$PDB = 1000 + X_1 + 2X_2 + 3X_3 + \varepsilon$$

Artinya adalah
$\beta_1$ = jika nilai variabel bebas lainya tetap konstan maka setiap kenaikan 1 unit investasi, PDB meningkat 1 unit
$\beta_2$ = jika nilai variabel bebas lainya tetap konstan maka setiap kenaikan 1 unit konsumsi, PDB meningkat 2 unit
$\beta_3$ = jika nilai variabel bebas lainya tetap konstan maka setiap kenaikan 1 unit ekspor, PDB meningkat 3 unit
$\beta_0$ = 1000 apabila tidak ada investasi, konsumsi, dan ekspor maka PDB masih bernilai 1000 unit.

e. Galat atau Kesalahan

Dalam regresi linear berganda, galat (*error*) atau kesalahan yang dilambangan dengan $\varepsilon$ adalah selisih antara nilai aktual variabel tetap ($Y$) dengan nilai prediksi yang dihasilkan oleh model regresi. Galat mencerminkan faktor-faktor lain yang tidak dimasukkan dalam model regresi atau ketidaksempurnaan model dalam menjelaskan hubungan antara variabel bebas dan variabel terikat.

Jenis-jenis galat dalam regresi linear berganda adalah galat acak dan galat sistematis. Galat acak (*Random Error*) adalah galat yang terjadi secara alami dan tidak bisa dihindari yang dikarenakan ada banyak faktor lain di luar

model yang mempengaruhi variabel terikat. Galat acak Biasanya dianggap tidak bermasalah apabila memenuhi asumsi klasik dalam regresi. Misalkan dalam bisnis, penjualan suatu produk dipengaruhi oleh permintaan konsumen dan harga dari produsen, tetapi ada faktor lain seperti cuaca yang menyebabkan variasi kecil dalam penjualan produk tersebut.

Galat sistematis (*Systematic Error*) merupakan galat yang terjadi karena ada pola atau bias dalam data yang tidak diperhitungkan dalam model regresi. Galat ini dapat mengarahkan pada hasil regresi yang salah. Misalkan dalam ekonomi: PDB suatu negara dipengaruhi oleh investasi dan konsumsi, tetapi jika model mengabaikan dampak besar dari kebijakan fiskal pemerintah, hasilnya bisa bias.

## 12.3 Estimasi Parameter Model

Estimasi parameter dalam regresi linear berganda bertujuan untuk menentukan koefisien regresi ($\beta_0$, $\beta_1$, ...,$\beta_n$) yang menggambarkan hubungan antara variabel bebas ($X$) dan variabel terikat ($Y$). Estimasi parameter dilakukan dengan metode *Ordinary Least Squares* (OLS) yang meminimalkan jumlah kuadrat galat (residual).

Cara kerja OLS dalam regresi linear berganda adalah sebagai berikut:

a. Menentukan Nilai Koefisien Regresi

Untuk menentukan koefisien regresi, dapat menggunakan rumus matriks. Rumus tersebut dapat digunakan untuk mendapatkan estimasi terbaik untuk setiap koefisien $\beta$ yang menghasilkan garis regresi optimal.

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

Keterangan:

$X$ = matriks variabel bebas (termasuk kolom konstanta $X_0$=1)

$Y$= matriks variabel terikat

$(X'X)^{-1}$= invers dari matriks $X'X$

$X'Y$= perkalian matriks transpos dari $X$ dan $Y$

b. Meminimalkan Jumlah Kuadrat Galat

OLS memilih garis regresi terbaik dengan meminimalkan Residual Sum of Squares (RSS):

$$RSS = \sum (Y_i - \widehat{Y}_i)^2$$

Keterangan:

$Y_i$ = nilai aktual dari variabel terikat

$\widehat{Y}_i$= nilai yang diprediksi oleh model regresi

Jika RSS kecil, maka model regresi lebih akurat dalam memprediksi nilai $Y$

**Contoh 1: Contoh Perhitungan OLS**

Sebuah perusahaan ingin mengetahui pengaruh biaya iklan ($X_1$) dan harga produk ($X_2$) terhadap penjualan ($Y$). Datanya sebagai berikut:

| Iklan ($X_1$) | Harga Produk ($X_2$) | Penjualan ($Y$) |
|---|---|---|
| 10 | 200 | 500 |
| 15 | 180 | 550 |
| 20 | 160 | 600 |

Tuliskan interprestasi koefisien regresinya!

Jawaban:

Dengan metode OLS, kita mendapatkan hasil estimasi

regresinya adalah Y= 0,285+16,662$X_1$+1,665$X_2$

Rinciannya sebagai berikut:

Persamaan regresi $Y = X\beta + \varepsilon$ bisa ditulis dalam bentuk matriks adalah

$$\begin{bmatrix}500\\550\\600\end{bmatrix} = \begin{bmatrix}1 & 10 & 200\\1 & 15 & 180\\1 & 20 & 160\end{bmatrix}\begin{bmatrix}\beta_0\\\beta_1\\\beta_2\end{bmatrix} + \begin{bmatrix}\varepsilon_0\\\varepsilon_1\\\varepsilon_2\end{bmatrix}$$

Menghitung koefisien regresi dengan rumus $\beta = (X'X)^{-1}X'Y$ sebagai berikut:

$$\beta = \begin{bmatrix}\beta_0\\\beta_1\\\beta_2\end{bmatrix}$$

$$= \left(\begin{bmatrix}1 & 1 & 1\\10 & 15 & 20\\200 & 180 & 160\end{bmatrix}\begin{bmatrix}1 & 10 & 200\\1 & 15 & 180\\1 & 20 & 160\end{bmatrix}\right)^{-1}\begin{bmatrix}1 & 1 & 1\\10 & 15 & 20\\200 & 180 & 160\end{bmatrix}\begin{bmatrix}500\\550\\600\end{bmatrix}$$

$$\beta = \begin{bmatrix}\beta_0\\\beta_1\\\beta_2\end{bmatrix} = \begin{bmatrix}0{,}285\\16.662\\1{,}665\end{bmatrix}$$

Interpretasi koefisien regresinya adalah

Interpretasi Koefisien:

1. Konstanta ($\beta_0$) = 0.285 artinya adalah saat biaya iklan ($X_1$) dan harga produk (X2) bernilai nol, maka penjualan diperkirakan 0.285. (Nilai ini biasanya tidak relevan secara praktis karena harga produk tidak mungkin bernilai nol)
2. Koefisien Biaya Iklan ($\beta_1$ )=16.662 artinya setiap penambahan 1 unit biaya iklan, maka penjualan meningkat sebesar 16.662 unit, dengan asumsi harga tetap.
3. Koefisien Harga Produk ($\beta_2$) = 1.665 artinya setiap kenaikan 1 unit harga produk, penjualan meningkat sebesar 1.665 unit, dengan asumsi biaya iklan tetap.

## 12.4 Uji Asumsi dalam Regresi Linear Berganda

a. Uji Linearitas

Uji linearitas bertujuan memastikan bahwa hubungan antara variabel bebas dan variabel terikat bersifat linear. Cara mengujinya adalah:

1. Scatter Plot (Diagram Pencar)

   Buat scatter plot antara masing-masing variabel independen dengan variabel dependen. Selanjutnya adalah memeriksa apakah titik-titik data mengikuti pola garis lurus atau membentuk kurva. Apabila titik-titik membentuk pola garis lurus , dapat dikatakan bahwa antar variabel memiliki hubungan linear sehingga asumsi terpenuhi. Apabila titik-titik membentuk pola melengkung (U atau lengkungan lainnya) maka dapat dikatakan bahwa hubungan antar variabek tidak linear sehingga asumsi tidak terpenuhi.

2. Uji Ramsey RESET (*Regression Specification Error Test*)

   Uji Ramsey RESET digunakan untuk memeriksa apakah ada hubungan non-linear yang tidak dimasukkan dalam model. Jika hasil uji menunjukkan bahwa model memiliki hubungan non-linear, maka kita mungkin perlu menggunakan transformasi variabel atau menambahkan variabel lain.

b. Uji Normalitas Residual

Uji normalitas residual bertujuan untuk memastikan bahwa residual (selisih antara nilai aktual dan nilai prediksi) berdistribusi normal. Normalitas residual diperlukan agar estimasi parameter dan uji statistik

dalam regresi valid, terutama dalam pengujian hipotesis dan pembuatan prediksi. Cara mengujinya adalah:

1. Histogram dan Q-Q Plot
   Histogram digunakan untuk membandingkan distribusi residual dengan distribusi normal. Sedangkan Q-Q Plot (*Quantile-Quantile Plot*) untuk mengetahui apabila titik-titik pada grafik Q-Q Plot membentuk garis lurus, maka residual berdistribusi normal.
   Jika histogram menunjukkan bentuk kurva lonceng (seperti distribusi normal), maka residual berdistribusi normal. Sedangkan apabila titik-titik pada Q-Q Plot berada di sekitar garis lurus, maka residual normal.
2. Uji Statistik Normalitas
   Uji statistik nomralitas dapat menggunakan Uji Kolmogorov-Smirnov (K-S Test), Uji Shapiro-Wilk, dan Uji Jarque-Bera.

c. Uji Homoskedastisitas (*Heteroskedastisitas*)
   Uji homoskedastisitas bertujuan memastikan bahwa varians residual tetap konstan di seluruh rentang nilai variabel independen. Jika tidak, terjadi heteroskedastisitas, yang dapat menyebabkan hasil regresi tidak efisien. Cara mengujinya adalah:
   1. Plot Scatter (Residual vs Nilai Prediksi)
      Buat scatter plot antara residual dan nilai prediksi. Jika titik-titik menyebar secara acak tanpa pola maka dapat dikatakan homoskedastisitas atau asumsi terpenuhi. Sedangan apabila titik-titik membentuk pola tertentu (misalnya mengecil atau membesar seperti kipas) maka dapat dikatakan heteros-kedastisitas atau asumsi tidak terpenuhi.

2. Uji Gletser
   Uji gletser berupa regresi antara nilai absolut residual dengan variabel independent. Jika nilai p-value > 0.05, maka tidak ada heteroskedastisitas.
3. Uji Breusch-Pagan
   Uji Breusch-Pagan menggunakan regresi tambahan untuk melihat apakah residual dipengaruhi oleh variabel independen. Jika p-value > 0.05, maka model dianggap homoskedastis.
4. Uji White
   Uji white mirip dengan Uji Breusch-Pagan, tetapi lebih fleksibel karena tidak mengasumsikan distribusi normal residual.

d. Uji Multikolinearitas
   Uji multikolinearitas bertujuan untuk memastikan bahwa variabel independen tidak memiliki hubungan yang sangat kuat satu sama lain. Jika ada multikolinearitas, maka interpretasi regresi bisa menjadi tidak valid. Jika terjadi multikolinearitas, maka koefisien regresi menjadi tidak stabil, signifikansi variabel bisa salah diinterpretasikan serta prediksi model menjadi tidak valid. Cara mengujinya adalah:
   a. *Variance Inflation Factor* (VIF)
      VIF mengukur seberapa besar varians dari koefisien regresi meningkat akibat adanya multikolinearitas. Rumus VIF untuk variabel:

$$VIF_j = \frac{1}{1 - R_j^2}$$

Keterangan:

$R_j^2$ adalah koefisien determinasi dari regresi variabel $X_j$ terhadap variabel independen lainnya.

Apabila VIF kurang dari 5 maka tidak ada multikolinearitas. Namun apabila nilai VIF antara 5 dengan 10 maka ada indikasi multikolinearitas ringan. Sedangkan jika nilai VIF lebih dari 10 maka dapat dikatakan nilai multikolinearitas tinggi dan model harus diperbaiki.

b. *Tolerance Value*

*Tolerance Value* adalah ukuran yang digunakan dalam uji multikolinearitas untuk mengetahui sejauh mana variabel independen tidak berkorelasi dengan variabel independen lainnya dalam model regresi linear berganda.

$$Tolerance\ 1 - R_j^2$$

Keterangan:

$R_j^2$ adalah koefisien determinasi dari regresi variabel $X_j$ terhadap variabel independen lainnya.

Apabila Tolerance > 0.1 maka dapat dikatakan bahwa tidak ada masalah multikolinearitas. Namun, apabila nilai *tolerance* lebih kecil dari 0.1 berarti ada masalah multikolinearitas.

c. Korelasi Antar Variabel Independen

Uji korelasi antar variabel independen berupa uji korelasi antara dua variabel. Jika nilai uji korelasi antar variabel independen lebih dari 0.8 atau lebih kecil dari negatif -0.8, maka ada indikasi multikolinearitas.

e. Uji Autokorelasi

Uji autokorelasi digunakan dalam analisis regresi untuk mengetahui apakah kesalahan residual (*error term*) memiliki pola tertentu dari satu observasi ke observasi lainnya. Apabila terjadi autokorelasi, maka asumsi regresi linear klasik tidak terpenuhi, yang dapat menyebabkan hasil regresi menjadi tidak akurat. Autokorelasi sering terjadi pada data deret waktu (*time series*), misalnya dalam analisis ekonomi, bisnis, dan keuangan.

## DAFTAR PUSTAKA

Efendi, Achmad, dkk. (2020). *Analisis Regresi: Teori dan Aplikasi dengan R*. Yogyakarta: ANDI.

Ghozali, Imam. (2018). *Aplikasi Analisis Multivariate dengan Program IBM SPSS 25*. Semarang: Badan Penerbit Universitas Diponegoro.

Gujarati, Damodar N. (2003). *Basic Econometrics (4th Edition).* New York: McGraw-Hill.

Maddala, G.S. (1992). *Introduction to Econometrics (2nd Edition).* New York: Macmillan Publishing Company.

Sugiyono. (2012). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.

Suyono. (2018). *Statistika untuk Penelitian*. Jakarta: Rajawali Pers.

# BAB 13
# BILANGAN INDEKS

**Oleh Fitriana Dina Rizkina**

## 13.1 Pengantar Bilangan Indeks

### 13.1.1 Definisi Bilangan Indeks

Pengertian bilangan indeks yaitu sebuah alat statistik yang berguna untuk mengukur perubahan relatif terhadap variabel tertentu, seperti total produksi, harga, atau tingkat inflasi pada suatu waktu tertentu, dibandingkan dengan waktu lainnya. Umumnya bilangan indeks didefinisikan sebagai rasio dari nilai data pada suatu période tertentu terhadap nilai data pada période dasar, dikalikan dengan 100 untuk mendapat nilai presentase. Contohnya, pada pengukuran inflasi, penggunaan indeks harga konsumen sering digunakan untuk menunjukkan perubahan harga barang dan jasa dari waktu ke waktu (Kaviza, 2019). Penggunaan bilangan indeks bertujuan untuk menyajikan gambar yang jelas mengenai trend dan fluktuasi data, yang memungkinkan untuk analisis lebih dalam dan informatif dalam konteks ekonomi dan bisnis.

Bilangan indeks menjadi penting dalam analisis statistik dan ekonomi karena kapasitasnya memberikan informasi dalam pengambilan keputusan dan pemahaman kondisi ekonomi. Melalui penggunaan

bilangan indeks, analisis ekonomi mampu mengukur kinerja ekonomi dari waktu ke waktu, menganalisis efek kebijakan ekonomi, dan memprediksi tren masa depan berdasarkan data historis (Achie, 2023). Pada era digital, bilangan indeks mampu menyajikan hasil analisis melalui variasi besar data ekonomi untuk interpretasi informasi kompleks dan pengambilan keputusan.

### 13.1.2 Sejarah dan Perkembangan Bilangan Indeks

Awal bilangan indeks diperkenalkan dapat dikaji kembali dengan penelusuran ke era awal statistik. Saat itu statistikawan dan ekonom mulai mencari metode pengukuran pada perubahan data ekonomi. Penggunaan bilangan indeks pertama kali pada awal abad ke-20, saat penerbitan laporan resmi berisi harga dan produksi oleh pemerintah (Curex, et al., 2022). Saat itu bilangan indeks muncul sebagai metode statistik yang digunakan untuk perhitungan perubahan harga relatif dari keranjang barang tertentu. Hal ini berkembang dan menjadi landasan bagi penggunaan bilangan indeks lebih lanjut dalam bidang analisis ekonomi.

Perkembangan teknologi dan metode analitis mempengaruhi revolusi penghitungan bilangan indeks. Ada beberapa pendekatan dalam perhitungan bilangan indeks, termasuk indeks Paasche dan indeks Fisher, menyajikan informasi akurat dalam kalkulasi. Penggunaan alat-alat statistik terus berinovasi, seperti regresi dan model matematis, telah mampu memberikan hasil lebih akurat dan relevan dalam penelitian dan analisis ekonomi (Sigalingging, et al., 2024). Melalui integrasi teknologi dan pendekatan statistik modern,

pemanfaatan bilangan indeks menjadi lebih fleksibel untuk data lebih kompleks dan komprehensif.

## 13.2 Tipe-Tipe Bilangan Indeks

### 13.2.1 Bilangan Indeks Harga

Pengertian bilangan indeks harga adalah sebuah alat matematis yang berguna untuk mengukur perubahan tingkat harga. Bilangan indeks harga memiliki peran dalam mencermati inflasi dan stabilitas ekonomi pada sebuah negara. Penggunaan dua bilangan indeks harga yang populer adalah indeks harga konsumen (CPI) dan indeks harga produsen (PPI).

Penggunaan indeks harga konsumen (CPI) adalah untuk mengukur perubahan rata-rata harga barang dan jasa (konsumsi rumah tangga). CPI berfungsi sebagai indikator inflasi dan penyajian informasi tentang dinamika daya beli konsumen (Prabuningrat et al., 2023). Aplikasi CPI juga mampu meramalkan dengan metode *Autoregressive Integrated Moving Averag*e (ARIMA), dimana hasil pendekatan ini menunjukkan cara efektif dalam memahami perubahan harga pada sebuah kota (Prabuningrat et al., 2023). CPI juga berperan untuk merumuskan kebijakan ekonomi pemerintah dalam pengontrolan inflasi dan pertumbuhan ekonomi (Salsabila dan Hasni, 2023). Selain itu, penggunaan indeks harga produsen (PPI) adalah untuk mengukur rata-rata dinamika harga yang diterima produsen domestik dalam usaha produk dan layanan mereka. Fungsi PPI adalah sebagai indikator awal dari dinamika harga dalam system ekonomi, yang nantinya akan mempengaruhi CPI. PPI berperan untuk

peramalan inflasi melalui perbandingan model peramalan.

Variasi metode penghitungan bilangan indeks tergantung penggunaan jenis indeks. Pada konteks CPI, misalnya, penggunaan metode penghitungan pemerintah berbasis pengambilan sampel dari harga barang-barang konsumen dan perubahan harga dari waktu ke waktu. Pendekatan seperti metode *double exponential smoothing* dan algoritma *support vector regression* berpotensi memberikan hasil keakuratan yang optimal dalam estimasi CPI (Septiawati et al., 2022; Suyono, et al., 2022; Ristyasari dan Ahdika, 2024). Hasil metode-metode ini memberikan informasi bahwa peramalan dengan MAPE (*mean absolute percentage error*) menunjukkan hasil analisis yang baik untuk nilai CPI di berbagai daerah, seperti Lampung dan Surabaya (Ristyasari & Ahdika, 2024; Septiawati et al., 2022).

### 13.2.2 Bilangan Indeks Kuantitas

Bilangan indeks kuantitas dapat didefinisikan sebagai salah satu metrik krusial pada analisis statistik dan ekonomi yang berguna memberi gambaran perubahan volume atau jumlah pada suatu kelompok data. Pada perhitungan dan pemodelan statistik, fungsi indeks untuk perbandingan nilai kuantitatif pada beberapa waktu yang berbeda atau konteks yang berbeda, melalui penggunaan metode matematis terstandarisasi. Ada sebuah pendekatan relevan pada konteks ini yaitu penggunaan indeks harga, yang biasanya digunakan dalam hasil penelitian. Pada sebuah penelitian, Rothwell indeks dinilai lebih rasional daripada Laspeyres indeks disebabkan kemampuannya

menunjukkan volatilitas harga dan kuantitas produksi, sebagai kunci pemahaman dinamika harga di pasar (Flukeria, 2022). Penelitian ini mendukung penggunaan indeks yang lebih responsif terhadap variabilitas harga, misalnya pada perumusan kebijakan pertanian. Selain itu, penelitian oleh Huda et al. (2022) menunjukkan wawasan tentang bagaimana pengukuran kuantitas pada analisis finansial, seperti penyusunan portofolio optimal terkait implementasi model indeks tunggal pada analisis kinerja investasi dan potensi risiko, menjadi semakin penting melalui penggunaan bilangan indeks kuantitas dalam pengambilan keputusan finansial yang lebih informatif (Huda et al., 2022)Secara komprehensif, penggunaan bilangan indeks kuantitas memiliki peran tidak hanya pada ekonomi dan statistik, namun juga alat penting dalam pengambilan keputusan lebih akurat di berbagai disiplin ilmu.

### 13.2.3 Bilangan Indeks Campuran

Penggunaan bilangan indeks campuran adalah untuk memperhitungkan kontribusi dari beberapa variabel, termasuk sifat fisik cairan, kondisi aliran, dan geometri alat, dalam pengukuran efisiensi proses pencampuran. Misalnya pada konteks mikromixer, implementasi bilangan indeks campuran menjadi penting dalam evaluasi performa pengadukan di level mikro, dimana terjadinya proses pencampuran di skala sangat kecil dan melibatkan fenomena fisik. Riset menunjukkan efisiensi pencampuran mikromixer sangat dipengaruhi oleh angka Reynolds (Re), sebagai rasio antara gaya viskos dan gaya inersia. Rentang angka Reynolds 0,1 hingga 500, mikromixer Kenics telah dibuktikan memiliki efisiensi pencampuran yang tinggi

dalam situasi yang mana indeks pencampuran dari fluida non-Newtonian menjadi fokus utama (Mahammedi et al., 2021). Hal ini disebabkan aliran fluida memiliki ciri berbeda pada angka Reynolds dimana berpengaruh pada efektivitas pencampuran. Hasil ini juga konsisten dengan riset Mahmud dan Tamrin (2020), yang menunjukkan pada Re rendah (<150), dominasi difusi molekular terjadi pada proses pencampuran. Sementara itu, pada Re yang lebih tinggi, waktu tinggal fluida lebih singkat sehingga indeks pencampuran menurun (Mahmud & Tamrin, 2020).

Geometri mikromixer memiliki pengaruh terhadap indeks pencampuran yang juga menjadi fokus dalam beberapa riset. Riset Jiao et al. (2022)melaporkan fungsi desain 3D pada struktur mikromixer mampu meningkatkan indeks pencampuran, melalui capaian hasil 99% efisiensi terhadap angka Reynolds di atas 20 (Jiao et al., 2022). Selain itu, pemodelan matematis juga menekankan pemahaman terhadap indeks pencampuran harus memperhatikan karakteristik perilaku fluida. Hal ini merupakan pengurangan indeks pencampuran yang terjadi dengan peningkatan nilai indeks power-law (Bordbar, et al., 2018). Maka, melalui optimalisasi geometri dan pemahaman karakter fluida dan non-Newtonian, peningkatan signifikan bisa dicapai pada efisiensi pencampuran. Secara umum, bilangan indeks campuran merupakan alat penting untuk memprediksi efisiensi pencampuran pada sistem mikrofusik. Peningkatan performa pencampuran bergantung pada pemahaman mendalam tentang faktor-faktor yang berpengaruh pada aliran dan interaksi molekular di skala mikro.

## 13.3 Metodologi Penghitungan Bilangan Indeks

### 13.3.1 Metode Statistika dalam Penghitungan

Salah satu aspek penting pada analisis data bisnis dan ekonomi adalah penghitungan bilangan indeks. Secara komprehensif, tujuan penggunaan bilangan indeks adalah untuk mengukur perubahan relatif dari beberapa variabel tertentu pada suatu waktu atau lokasi. Metode penghitungan bilangan indeks yang sering digunakan mencakup Indeks Harga Konsumen (IHK) dan Indeks Produksi, yang mana dibagi menjadi dua kategori yaitu Laspeyres dan Paasche. Pada metode Laspeyres, digunakan tahun dasar dalam pengukuran perubahan, bahwa harga dan kuantitas pada tahun dasar diringkas untuk période perbandingan harga dan kuantitas. Sementara itu, pada metode Paasche, diperhitungkan kuantitas pada tahun berjalan, sehingga mampu menyajikan gambaran lebih fleksibel tentang harga dan kuantitas saat ini (Almasah & Sirait, 2024). Metode statistik juga digunakan secara lebih luas yang melibatkan analisis varians dan penggunaan regresi dalam mengontrol variabel lain yang mungkin mempengaruhi hasil (Rachmawati, et al., 2023). Analisis ini juga membutuhkan penanganan data yang tepat, seperti ketika berhadapan dengan data time series, maka pemilihan alat statistik yang tepat dan metodologis kuat untuk mendapatkan hasil yang *reliable* (dapat dipercaya) dan valid (Fauzan, et al., 2020).

### 13.3.2 Program dan Alat Bantu

Investasi pada perangkat lunak statistik atau alat bantu analisis menjadi sangat penting pada perhitungan bilangan indeks. Beberapa perangkat lunak yang

populer digunakan seperti SPSS, Python dan R, yang memiliki fitur-fitu sangat baik dalam analisis data, visualisasi dan modeling (Friyana, et al., 2023). Penggunaan R semakin banyak digunakan karena dukungan ketersediaan pustaka yang kaya dalam analisis statistik dan bisa digunakan dalam menghitung berbagai jenis indeks dengan metode analisis berbeda (Jambormias et al., 2014). Selain itu, penggunaan alat bantu lain, termasuk alat digital dan platform berbasis cloud, dapat meningkatkan pengguna dalam mengakses dan memanipulasi data berkapasitas besar dengan efisien dan efektif. Contohnya, fitur pada Microsoft Excel dapat digunakan menghitung berbagai indeks dengan relatif mudah, terutama kapasitas data kecil sampai menengah. Di samping itu, data satelit dan model ekonometrika juga dapat dimanfaatkan dalam meningkatkan akurasi dan efektivitas penghitungan indeks, khususnya pada konteks ekonomi makro (Simangunsong, et al, 2022). Secara komprehensif, penggunaan alat bantu dan metode statistika yang tepat dapat memberikan hasil analisis bilangan indeks yang kuat sebagai fondasi pengambilan keputusan lebih valid dan analitis.

## 13.4 Aplikasi Praktis dari Bilangan Indeks

### 13.4.1 Analisis Ekonomi Makro

Bilangan indeks memiliki peran vital pada analisis ekonomi makro yang membantu dalam pengukuran dan pemantauan indikator ekonomi. Salah satu implementasi bilangan indeks yaitu penggunaan dalam pengukuran inflasi melalui Indeks Harga Konsumen

(IHK). Data harga barang dan jasa mampu dilakukan analisis melalui IHK untuk pengamatan dinamika daya beli masyarakat (Santoso dan Rakhmawan, 2021). Bilangan indeks juga penting untuk menganalisis tren inflasi dan adaptasi kebijakan moneter untuk mengatasi terjadinya fluktuasi di pasar.

Penggunaan bilangan indeks juga digunakan dalam pengukuran kinerja sektor ekonomi tertentu, seperti produksi industri (Sudarsono, 2018). Melalui penggunaan bilangan indeks produksi, praktisi dan pengamat ekonomi mampu mengevaluasi pertumbuhan atau penurunan produksi. Selain itu, ada hubungan signifikan antara Indeks Saham Syariah Indonesia (ISSI) dan indikator-indikator makroekonomi, seperti suku bunga dan indeks harga (Sudarsono, 2018). Maka dari itu, bilangan indeks mampu memberikan perspektif komprehensif terhadap kondisi ekonomi makro.

### 13.4.2 Penerapan di Dunia Bisnis

Pada konteks dunia bisnis, penggunaan bilangan indeks diaplikasikan untuk evaluasi kinerja perusahaan dan pengukuran daya saing di pasar. Misalnya, Indeks Keberlanjutan, kini semakin dipertimbangkan oleh investor untuk mengevaluasi kinerja perusahaan khususnya dalam sosial dan lingkungan (Uljannah dan Ibrahim, 2023). Kemudian, Indeks Harga Saham, metode yang digunakan untuk analisis kinerja saham di Bursa Efek Indonesia (BEI). Fluktuasi indeks harga saham gabungan seringkali dipengaruhi oleh faktor-faktor ekonomi, seperti perubahan kondisi bisnis di Indonesia (Handika, et al., 2021). Di samping itu, penggunaan bilangan indeks juga digunakan dalam pengukuran dampak dari sebuah kebijakan (Hofifah, 2020),

contohnya pemberian insentif atau pengurangan pajak terhadap kinerja perusahaan dan pasar.

### 13.4.3 Penggunaan dalam Penelitian dan Kebijakan Publik

Pada konteks kebijakan publik dan penelitian, penggunaan bilangan indeks diaplikasikan dalam penyediaan data sebagai pertanggungjawaban pembuat kebijakan. Penggunaan Indeks Pekerjaan Layak (IPL) berperan menyajikan informasi terhadap penyusunan kebijakan lebih efisien dan efektif, seperti pada isu ketenagakerjaan (Santoso dan Rakhmawan, 2021). Analisis IPL selama masa pandemi telah berhasil mengidentifikasi rekomendasi strategi dalam perbaikan kualitas kebijakan pekerjaan di Indonesia (Santoso dan Rakhmawan, 2021). Fungsi bilangan indeks juga menjadi alat evaluasi dalam kebijakan publik yang terkait kesehatan, lingkungan, dan pendidikan. Contohnya, sistem pengukuran emisi gas rumah kaca atau kualitas udara (Kautsar, et al., 2024).

## 13.5 Tantangan dan Pertimbangan dalam Penggunaan Bilangan Indeks

### 13.5.1 Keterbatasan Metodologi

Pada analisis statistik dan ekonomi, penggunaan bilangan indeks berhadapan dengan berbagai keterbatasan metodologi. Pertama, bilangan indeks bergantung pada ketersediaan data, dimana sebagai gambaran informasi kondisi sebenarnya. Misalnya kelengkapan dan kualitas data memiliki pengaruh terhadap keakuratan indeks. Saat data yang tidak

representatif digunakan, hasil analisis dapat memberikan informasi yang salah bagi pengambilan keputusan (Widyatmoko dan Setiawan, 2023).

Kedua, kemungkinan bias dapat terjadi dalam seleksi indikator dalam perhitungan bilangan indeks. Beberapa variabel tertentu yang digunakan mampu menyajikan kondisi ekonomi lebih baik daripada yang lain. Perhitungan ini berpotensi menghasilkan indeks yang tidak komprehensif dan tidak informatif. Riset menunjukkan adanya risiko pada pengkategorian wilayah berdasarkan potensi ekonomi, mempengaruhi perkembangan usulan kebijakan (Prasetyo, 2022). Di samping itu, variasi metode mampu menghasilkan variasi hasi yang signifikan walaupun data yang digunakan sama, maka pemilihan metode yang relevan menjadi penting dilakukan (Gayatri et al., 2023).

### 13.5.2 Implikasi Etis dan Kebijakan

Pertimbangan etis dan kebijakan sering kali mempengaruhi penggunaan bilangan indeks. Hal ini dapat dicermati pada penyusunan kebijakan publik dimana bilangan indeks perlu diberikan pemahaman dan penjelasan yang jelas kepada publik sehingga dapat dipertanggungjawabkan. Penyampaian informasi ini akan memberikan interpretasi kuat pada masyarakat dalam kepercayaan pada proses perumusan kebijakan (Saleh et al., 2022). Penggunaan bilangan indeks sering diaplikasikan dalam pengambilan keputusan/kebijakan dan pertimbangan etis. Contohnya, pada konteks proses penyusunan kebijakan publik, penjelasan dan pemahaman bilangan indeks perlu dipaparkan dengan jelas dan valid kepada publik. Apabila penyampaian informasi disampaikan tidak jelas, maka akan

berpeluang merugikan pihak tertentu atau berkontribusi pada keputusan tidak tepat (Saleh et al., 2022).

Pemilihan indikator yang beretika perlu diperhatikan dalam penggunaan bilangan indeks. Indikator yang dipilih nantinya perlu disusun dengan cermat untuk menghindari pola diskriminasi atau stigma pada kelompok tertentu, misalnya pada isu-isu sensitif, seperti ketidaksetaraan dan kemiskinan (Tandiawan, 2022). Sebuah riset melaporkan bahwa pertimbangan aspek etik menjadi penting dalam penyusunan metodologi survei sehingga tidak menimbulkan kesalahan informasi atau ketidakpuasan masyarakat (Paulangan, et al., 2021). Maka dari itu, implementasi bilangan indeks menjadi sangat dipertimbangkan terutama dampaknya dalam teknis, sosial dan etis.

## 13.6 Kesimpulan dan Rekomendasi

Bilangan indeks pada konteks bisnis dan ekonomi menunjukkan bahwa bilangan indeks adalah alat analisis signifikan yang dapat digunakan pada pengukuran perubahan ekonomi dan pengambilan keputusan berbasis data. Pembuat kebijakan dan peneliti dapat lebih memahami dinamika ekonomi berdasarkan pengukuran variabel-variabel. Implementasi bilangan indeks telah terbukti menyajikan wawasan mendalam terhadap output dan perubahan harga, seperti Indeks Harga Konsumen dan Indeks Produksi (Widyatmoko dan Setiawan, 2023). Di samping itu, pemahaman mengenai fluktuasi indikator ekonomi berpotensi mendukung manajer untuk menyusun

strategi bisnis, seperti pada reaksi pasar terhadap perubahan indeks tertentu (Widyatmoko dan Setiawan, 2023).

Ada tantangan pada implementasi bilangan indeks yang berdampak pada akurasi hasil. Keterbatasan bilangan indeks termasuk kualitas data, yang berpotensi tidak konsisten dan pemilihan indikator yang mungkin bias dapat merugikan proses pengambilan keputusan. Maka dari itu, evaluasi terhadap metodologi dalam perhitungan bilangan indeks menjadi penting (Wahyuningtyas, 2015). Penggunaan bilangan indeks juga perlu mempertimbangkan etika seperti memperhatikan implikasi sosial dari kebijakan yang akan dihasilkan. Pada riset kemiskinan, penting menggunakan indeks yang mempertimbangkan aspek ekonomi dan sosial (Gopal, et al., 2023). Secara komprehensif, bilangan indeks akan menjadi alat statistik dan instrumen penting pada pemahaman perubahan ekonomi dan sosial.

## DAFTAR PUSTAKA

Achie, N. (2023). Wacana Diskursif Dalam Pensejarahan Islam: Penilaian Kritis Terhadap Pemikiran Mohd. Asri Zainul Abidin Dalam Pertelingkahan Para Sahabat Nabi – Antara Ketulenan Fakta Dengan Pembohongan Sejarah. *Jurnal Kinabalu*, *29*, 1–22. https://doi.org/10.51200/ejk.v29i.4782

Almasah, M. Z., & Sirait, T. (2024). Analisis Ekonomi Sosial Dan Indeks Inklusi Keuangan Di Indonesia Dengan Persamaan Bentuk Tereduksi. *Limits Journal of Mathematics and Its Applications*, *21*(1), 81. https://doi.org/10.12962/limits.v21i1.17113

Bordbar, A., Taassob, A., & Kamali, R. (2018). Diffusion and Convection Mixing of Non-Newtonian Liquids in an Optimized Micromixer. *The Canadian Journal of Chemical Engineering*, *96*(8), 1829–1836. https://doi.org/10.1002/cjce.23113

Curex, N., Toaha, S., & Kasbawati, K. (2022). Stability Analysis of the Development Mathematical Models for the Spread of Covid-19 Effects of Vaccination and Campaigns as Processes in Controlling the Spread of Disease. *Proximal Jurnal Penelitian Matematika Dan Pendidikan Matematika*, *5*(2), 50–66. https://doi.org/10.30605/proximal.v5i2.1819

Fauzan, I. F., Firdaus, M., & Sahara, S. (2020). Regional Financial Inclusion and Poverty: Evidence From Indonesia. *Economic Journal of Emerging Markets*, *12*(1), 25–38. https://doi.org/10.20885/ejem.vol12.iss1.art3

Flukeria, M. (2022). Perlukah Alternatif Penghitungan Nilai Tukar Petani? Simulasi Perbandingan Indeks Harga Laspeyres Index Dan Rothwell Index Pada Komoditas Ikan Segar Di Indonesia. *Seminar Nasional Official Statistics*, *2022*(1), 79–90. https://doi.org/10.34123/semnasoffstat.v2022i1.1160

Friyana, A. O., Harisuseso, D., & Andawayanti, U. (2023). Pemanfaatan Data Satelit Untuk Menganalisis Indeks Kekeringan Meteorologi Di Sub DAS Slahung Kabupaten Ponorogo. *Jurnal Teknologi Dan Rekayasa Sumber Daya Air*, *4*(1), 291–301. https://doi.org/10.21776/ub.jtresda.2024.004.01.024

Gayatri, M. R., Fadhilah, R., Lestari, S. T., Pratiwi, L. F., Abdurahim, A., & Wardhana, I. G. A. W. (2023). Topology Index of the Coprime Graph for Dihedral Group of Prime Power Order. *Jurnal Diferensial*, *5*(2), 126–134. https://doi.org/10.35508/jd.v5i2.12462

Gopal, P. S., Malek, N. M., & Rahman, M. A. N. A. (2023). Penilaian Pengukuran Kemiskinan Dari Perspektif Multidimensi: Kajian Kes Pelajar B40 Di Universiti Sains Malaysia. *Malaysian Journal of Social Sciences and Humanities (Mjssh)*, *8*(12), e002616. https://doi.org/10.47405/mjssh.v8i12.2616

Handika, H., Damajanti, A., & Rosyati, R. (2021). Faktor Penentu Fluktuasi Pergerakan Indeks Harga Saham Gabungan (Ihsg) Di Bursa Efek Indonesia (Bei). *Solusi*, *19*(3), 291. https://doi.org/10.26623/slsi.v19i3.3503

Hofifah, S. (2020). Analisis Persaingan Usaha Pedagang Musiman Di Ngebel Ponorogo Ditinjau Dari Perspektif Etika Bisnis Islam. *Syarikat Jurnal Rumpun Ekonomi*

*Syariah*, *3*(2), 37–44. https://doi.org/10.25299/syarikat.2020.vol3(2).6469

Huda, M., Abdullah, S., Chasanah, S. I. U., Mursyidah, H., Ikhsan, F., Susilo, S., & Sukandar, R. S. (2022). Analisis Pembentukan Portofolio Optimal Saham-Saham JII30 Dengan Model Indeks Tunggal Periode New-Normal. *Jurnal Derivat Jurnal Matematika Dan Pendidikan Matematika*, *9*(1), 32–46. https://doi.org/10.31316/j.derivat.v9i1.2758

Jambormias, E., Sutjahjo, S. H., Mattjik, A. A., Wahyu, Y., & Wirnas, D. (2014). Perluasan Indeks Seleksi Nilai Fenotipe Untuk Indeks Seleksi Nilai Pemuliaan. *Buletin Agrohorti*, *2*(1), 115. https://doi.org/10.29244/agrob.2.1.115-124

Jiao, D., Zhang, R., Zhang, H., Ren, S., Feng, H., & Chang, H. (2022). Performance Evaluation of a 3D Split-and-Recombination Micromixer With Asymmetric Structures. *Journal of Micromechanics and Microengineering*, *32*(7), 75007. https://doi.org/10.1088/1361-6439/ac7771

Kautsar, S., Ridwansyah, R. R., & Priscilla, N. (2024). Pembentukan Indeks Emisi Gas Rumah Kaca Dan Strategi Menuju Net Zero Emission Kabupaten/Kota Di Pulau Jawa Dalam Mendukung Indonesia Emas 2045. *Seminar Nasional Official Statistics*, *2024*(1), 591–602. https://doi.org/10.34123/semnasoffstat.v2024i1.2119

Kaviza, M. (2019). Korelasi Antara Motivasi Intrinsik Dengan Pencapaian Subjek Sejarah. *Malaysian Journal of Social Sciences and Humanities (Mjssh)*, *4*(8), 1–10. https://doi.org/10.47405/mjssh.v4i8.327

Mahammedi, A., Tayeb, N. T., Kim, K., & Hossain, S. (2021). Mixing Enhancement of Non-Newtonian Shear-Thinning Fluid for a Kenics Micromixer. *Micromachines*, *12*(12), 1494. https://doi.org/10.3390/mi12121494

Mahmud, F., & Tamrin, K. F. (2020). Method for Determining Mixing Index in Microfluidics by RGB Color Model. *Asia-Pacific Journal of Chemical Engineering*, *15*(2). https://doi.org/10.1002/apj.2407

Paulangan, Y. P., Supoyo, A. S., & Kalor, J. D. (2021). Density and Ecological Index of Nudibranch in Humbolt Bay Water Jayapura City Papua Province, Indonesia. *Tropical Fisheries Management Journal*, *5*(1), 59–64. https://doi.org/10.29244/jppt.v5i1.34406

Prabuningrat, S. H., Haris, M. Al, Salma, N. K., Muharamah, P. W., & Nur, M. S. (2023). Peramalan Indeks Harga Konsumen Kota Semarang Dengan Metode Autoregressive Integrated Moving Average. *Journal of Data Insights*, *1*(1), 1–9. https://doi.org/10.26714/jodi.v1i1.124

Prasetyo, H. (2022). Regional Clustering Based on Economic Potential With a Modified Fuzzy K-Prototypes Algorithm for Village Developing Target Determination. *Jurnal Teknologi Dan Sistem Komputer*, *10*(1), 46–52. https://doi.org/10.14710/jtsiskom.2021.14247

Rachmawati, F., Sigalingging, N. Y., & Kiftiah, M. (2023). Proyeksi Indeks Pembangunan Gender (IPG) Di Kalimantan Barat Dengan Metode Trend Parabolik. *Jurnal Forum Analisis Statistik (Formasi)*, *2*(2), 83–91. https://doi.org/10.57059/formasi.v2i2.33

Ristyasari, D. P., & Ahdika, A. (2024). *Peramalan Indeks Harga Konsumen Di Kota Bandar Lampung Tahun 2023*

*Menggunakan Pemodelan Double Exponential Smoothing*. *2*(1), 145–152. https://doi.org/10.20885/esds.vol2.iss.1.art14

Saleh, S. F., Nasrun, N., Sulfasyah, S., Damayanti, A., Nurwahida, N., & Isnayanti, A. N. (2022). Pelatihan Media Pembelajaran Operasi Hitung Bilangan Bulat Bagi Calon Guru Sekolah Dasar. *Bima Abdi Jurnal Pengabdian Masyarakat*, *2*(2), 47–55. https://doi.org/10.53299/bajpm.v2i2.196

Salsabila, Q., & Hasni, N. I. (2023). Analisis Pengaruh Indeks Harga Konsumen Bulanan Terhadap Inflasi Makanan Bulanan Melalui Metode Analisis Regresis Sederhana. *Cidea Journal*, *2*(2), 111–116. https://doi.org/10.56444/cideajournal.v2i2.1311

Santoso, K. N., & Rakhmawan, S. A. (2021). Indeks Komposit Pekerjaan Layak Di Indonesia Pada Era Pandemi COVID-19. *Seminar Nasional Official Statistics*, *2021*(1), 214–222. https://doi.org/10.34123/semnasoffstat.v2021i1.840

Septiawati, D., Gusti, S. K., Syafria, F., Yusra, Y., & Cynthia, E. P. (2022). Prediksi Data Indeks Harga Konsumen Provinsi Riau Berbasis Time Series Dengan Metode Double Exponential Smoothing. *Jipi (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, *7*(4), 1342–1350. https://doi.org/10.29100/jipi.v7i4.3209

Sigalingging, N. Y., Fran, F., & Kusumastuti, N. (2024). Menentukan Invers Drazin Dengan Teorema Cayley Hamilton. *Mathvision Jurnal Matematika*, *6*(1), 11–16. https://doi.org/10.55719/mv.v6i1.998

Simangunsong, A., Simanjorang, R. M., & Fahmi, H. (2022). *Penerapan Metode Composite Performance Index Dalam*

*Seleksi Penerimaan Calon Laboran*. *1*(2), 41–48. https://doi.org/10.55338/justikpen.v1i2.8

Sudarsono, H. (2018). Indikator Makroekonomi Dan Pengaruhnya Terhadap Indeks Saham Syariah Di Indonesia. *Esensi Jurnal Bisnis Dan Manajemen*, *8*(2). https://doi.org/10.15408/ess.v8i2.7219

Suyono, A. A., Kusrini, K., & Arief, M. R. (2022). Prediksi Indeks Harga Konsumen Komoditas Makanan Di Kota Surabaya Menggunakan Support Vector Regression. *Metik Jurnal*, *6*(1), 45–51. https://doi.org/10.47002/metik.v6i1.339

Tandiawan, W. (2022). Tanggung Jawab Sosial Perusahaan Dan Akses Pendanaan. *Jurnal Akuntansi Keuangan Dan Manajemen*, *3*(4), 323–341. https://doi.org/10.35912/jakman.v3i4.1392

Uljannah, A., & Ibrahim, H. (2023). Analisis Etika Dalam Konteks Lintas Budaya Dalam Bisnis Internasional. *Jurnal Minfo Polgan*, *12*(2), 2438–2443. https://doi.org/10.33395/jmp.v12i2.13295

Wahyuningtyas, D. T. (2015). Penggunaan Media Mobil Mainan Untuk Meningkatkan Pemahaman Konsep Operasi Hitung Bilangan Bulat. *Jurnal Inspirasi Pendidikan*, *5*(1), 587. https://doi.org/10.21067/jip.v5i1.689

Widyatmoko, A. S., & Setiawan, D. (2023). Reaksi Pasar Pada Peristiwa Rebalancing Indeks Msci Di Negara Asean-5. *Jurnal Kajian Akuntansi*, *7*(1), 35. https://doi.org/10.33603/jka.v7i1.7217

## BIODATA PENULIS

**Pelliyezer Karo Karo, ST, MM, CHE, CIQnR, CPM (Asia), CRA®, CRP®**
Dosen Politeknik Pariwisata Palembang
Kementerian Pariwisata RI

Penulis menyelesaikan studi S1 di Program Studi Teknik Elektro Fakultas Teknik Universitas Sumatera Utara, kemudian melanjutkan studi S2 pada Program Magister Manajemen Sekolah Pascasarjana Universitas Sumatera Utara. Saat ini sedang aktif menempuh studi S3 pada Program Doktor Ilmu Manajemen Fakultas Ekonomi Universitas Sriwijaya. Penulis merupakan dosen tetap Politeknik Pariwisata Palembang, Kementerian Pariwisata Republik Indonesia. Menjalankan darma pelaksanaan pendidikan dengan bidang ilmu utama Manajemen Pemasaran. Selain itu, mata kuliah yang juga menjadi area keahlian adalah Metodologi Penelitian dan Statistika. Dalam memenuhi kewajiban Tridarma, penulis tertarik untuk berkonsentrasi pada darma penelitian sebagai prioritas utama dengan aktif menghasilkan publikasi ilmiah secara konsisten.

## BIODATA PENULIS

**Anita Roosmawarni, SE., M.SE.**
Dosen Program Studi Manajemen
FEB Universitas Muhammadiyah Surabaya

Lulus S1 pada Program Studi Ilmu Ekonomi dan Studi Pembangunan, Fakultas Ekonomi dan Bisnis Universitas Brawijaya, Malang tahun 2000, lulus S2 di Program Magister Ilmu Ekonomi Universitas Airlangga tahun 2014, dan tengah menempuh Program Doktoral Ilmu Ekonomi Universitas Airlangga, Surabaya. Saat ini adalah dosen tetap Program Studi Manajemen Universitas Muhammadiyah Surabaya. Pernah menjadi dosen tamu dalam Program Guest Lecture di Universiti Selangor (Unisel) Malaysia dengan tema Marketing Framework of Small and Medium Enterprise (SMEs) Based on Marketing Mix of Syariah. Aktif melakukan pendampingan pada kegiatan abdimas seperti KKN KP KAS Tahun 2018 dan Tahun 2020. Pernah memperoleh Hibah Ristek Dikti di Tahun 2017 melalui dua skema, yaitu PUPT dan PDP serta melahirkan beberapa buku seperti : Buku Ajar Marketing Syariah dan Buku Kewirausahaan (Dasar & Konsep)

## BIODATA PENULIS

**Dr.Ir. Patrich Phill Edrich Papilaya. M.Sc.Forest.Trop.**
Dosen Program Studi Pascasarjana Manajemen Hutan
Universitas Pattimura

Penulis lahir di Ambon Tanggal 14 Pebruari 1967. Penulis adalah dosen tetap pada Program Studi Manajemen Hutan Pascasarjana. Universitas Pattimura. Ambon. Menyelesaikan pendidikan S1 pada Jurusan Kehutanan Fakultas Pertanian Universitas Pattimura. menyelesaikan Pendidikan S2 pada Georg-August Universitas. Faculty of Forestry. Göttingen. dalam bidang Remote sensing dan GIS. Pendidikan S3 diselesaikan pada IPB University. Bogor

## BIODATA PENULIS

**Toto Hermawan S.Pd.,M.Sc.**
Dosen Program Studi Pendidikan Matematika
Fakultas Keguruan dan Ilmu Pendidikan

Penulis lahir di Bantul tanggal 2 desember 1988. Penulis adalah dosen tetap pada Program Studi Pendidikan Matematika Fakultas Keguruan dan Ilmu Pendidikan, Universitas Cokroaminoto Yogyakarta. Menyelesaikan pendidikan S1 Pendidikan Matematika Universitas Ahmad Dahlan ( 2011), S2 Pada Program studi Magister Matematika Universitas Gajah Mada ( 2015). Penulis menekuni bidang Penelitian dan Menulis buku ajar. Selain itu Penulis mengampu mata kuliah Pengelolaan Data Statistik, Statitika Matematika , Teori Bilangan, Teori Probabilitas dan Statitika Penelitian

## BIODATA PENULIS

**Muhammad Zulkarnain, ME**
Dosen Fakultas Ekonomi dan Bisnis Islam
UIN Sjech M. Djamil Djambek Bukittinggi

Penulis adalah putra pertama (dari sembilan bersaudara) dari Bapak Abdul Halim (*Almarhum*) dan Ibu Asro yang dilahirkan di Asahan (Sumatera Utara) pada 10 Februari 1967. Setelah menamatkan pendidikan SD tahun 1980, SMP tahun 1983, dan SMA tahun 1986, kemudian melanjutkan pendidikan Sarjana ($S_1$) di Jurusan Sosial Ekonomi Pertanian (*Agribisnis*) Fakultas Pertanian Institut Pertanian "STIPER" (INSTIPER) Yogyakarta, dan lulus pada tahun 1994 dengan judul Tesis *Evaluasi Perkreditan Petani Plasma Kelapa Sawit di PIR Khusus I Sungei Tapung Riau*.

Pada tahun 1995-2000, penulis bekerja di Hutan Tanaman Industri (HTI) PT. PSPI Surya Dumai Group Pekanbaru. Lalu berwiraswasta mulai tahun 2001, kemudian pada bulan Juli 2017, melanjutkan pendidikan $S_2$ pada Program Studi Magister Ekonomi Syariah Fakultas Ekonomi dan Bisnis Islam (FEBI) Institut Agama Islam Negeri (IAIN) Bukittinggi, dan lulus pada bulan Agustus 2019 dengan judul Tesis Pengaruh *Financing to Deposit Ratio* (FDR) dan *Non Performing Financing* (NPF) terhadap Profitabilitas Bank

Pembiayaan Rakyat Syariah di Sumatera Barat dengan Inflasi sebagai Variabel Moderasi.

Karya tulis yang telah terbit berupa buku-buku yang diterbitkan oleh Penerbit CV. Media Sains Indonesia sebagai Penulis Book Chapter, berjudul *Ilmu Komunikasi & Statistik*, dan buku-buku yang diterbitkan oleh Penerbit Get Press Indonesia sebagai Penulis Book Chapter, berjudul *Ekonomi Pembangunan Islam, Model Ekonomi Syariah Fondasi Sistem Ekonomi, Sistem Ekonomi Internasional, Ekonomi Mikro Islam 2, Statistik Ekonomi dan Bisnis: Konsep dan Metode Analisis*.

Artikel Jurnal yang telah terbit berjudul *Peran Non Performing Financing terhadap Profitabilitas Bank Pembiayaan Rakyat Syariah di Sumatera Barat dengan Inflasi sebagai Variabel Moderasi* yang diterbitkan oleh Ekonomika Syariah Journal of Economic Studies, dengan link : *https://ejournal.uinbukittinggi.ac.id/index.php/febi/article/view/3305.*

Selanjutnya, sampai dengan sekarang penulis aktif sebagai Dosen mata kuliah *Komunikasi Bisnis Islam, Ekonomi Internasional, Ekonomi Mikro Syariah, Ekonomi Makro Syariah, Fiskal dan Moneter dalam Islam, Pengantar Kewirausahaan, Ekonomi Publik, Etika Bisnis Syariah, Asuransi Syariah, Matematika Ekonomi, dan Statistik Ekonomi* di Fakultas Ekonomi dan Bisnis Islam (FEBI) Universitas Islam Negeri (UIN) Sjech M. Djamil Djambek Bukittinggi.

## BIODATA PENULIS

**Ir. Roosganda Elizabeth, M.Si., MS.PNLP.**
BRIN

Penulis, berkarya nyata sebagai fungsionalis peneliti ASN – Badan Riset dan Inovasi Nasional (BRIN) di bidang Pemberdayaan dan Inovasi Kompetensi SDM, Kelembagaan, Sosial Ekonomi Manajemen Bisnis dan Agribisnis, Ekonomi Khusus, Kerakyatan dan Kreatif, Pengabdian Masyarakat, Founder Start Up, Dewan Komisaris Perkebunan Swasta Nasional, Mitra Bestari, Reviewer, dan Editor anggota Tim Dewan Redaksi dan Editorial Board beberapa publikasi dan jurnal.

Aktif menulis karya tulis ilmiah (KTI) di berbagai publikasi nasional dan internasiona dengan 400 an artikel telah publis di buku, jurnal internasional dan nasional dan memperoleh penghargaan lebih dari 75 HKI. Penulis berkarya nyata sebagai fungsionalis dan pengurus aktif di beberapa organisasi dan profesi ilmiah. Ber-orientasi sektor publik dengan menyelaraskan keberpihakan dan pendampingan. Pembimbing Mahasiswa agar mampu menyelesaikan tugas akhir, skripsi, tesis, desertasi

berkualitas dan sempurna agar segera mampu memanfaatkan keilmuannya di masyarakat dan lingkungan luas dengan kompetensi, qualified, bernilai jual tinggi dan bertanggungjawab serta ber-ahlak, ber-attitude, menjunjung tinggi dan komit melaksanakan etika dan norma empati, simpati dan toleransi dalam bermasyarakat yang majemuk dan beragam, untuk dan supaya mampu menuliskan, mendiseminasikan, mensosialisasikan dan menyebarluaskan hasil karya nyata mereka kepada masyarakat luas (nasional dan internasional) melalui karya tulis ilmiah, artikel, deskripsi dan statement ilmiah dan populer yang kompeten, bertanggungjawab, bermanfaat tinggi dan luas.

## BIODATA PENULIS

**Joni Wilson Sitopu**
Dosen Universitas Simalungun

Joni Wilson Sitopu, lahir di Medan pada tanggal 25 Mei 1964, Pendidikan S1 Prodi Matematika FMIPA USU Medan, S2 Prodi Pendidikan Matematika UNIMED Medan dan S3 Prodi Ilmu Matematika FMIPA USU Medan. Dosen di Teknik Sipil, SPS Prodi S2 PW, IPS dan Manajemen, serta Kepala Lembaga Akreditasi Internal (LAI) Universitas Simalungun. Penulis aktif mengikuti sebagai pemakalah pada Seminar Nasional dan internasional, aktif mengikuti Seminar Nasional dan internasional dan aktif menulis Buku, Jurnal OJS, Sinta, dan Scopus.

## BIODATA PENULIS

**Nadia Fazira, S.E., M.Sc**
Dosen Program Studi Ekonomi Pembangunan
Fakultas Ekonomi dan Bisnis Universitas Andalas

Penulis lahir di Bukittinggi pada tanggal 5 Desember 1994. Saat ini, penulis merupakan dosen tetap pada Program Studi Ekonomi Pembangunan, Fakultas Ekonomi dan Bisnis, Universitas Andalas. Gelar Sarjana Ekonomi diperoleh dari Universitas Andalas tahun 2016, dan gelar Magister Ilmu Ekonomi dari Universitas Gadjah Mada tahun 2019.

## BIODATA PENULIS

**Asyidatur Rosmaniar,SE., M.Pd.**
Dosen Program Studi Manajemen
Fakultas Ekonomi dan Bisnis
Universitas Muhammadiyah Surabaya

Penulis lahir di Surabaya tanggal 12 Maret 1981. Penulis adalah dosen tetap pada Program Studi Program Studi Manajemen Fakultas Ekonomi dan Bisnis Universitas Muhammadiyah Surabaya. Menyelesaikan pendidikan S1 pada Jurusan Manajemen Fakultas Ekonomi Universitas Airlangga dan melanjutkan S2 pada Jurusan Manajemen Pendidikan Universitas Negeri Surabaya. Penulis mengampu mata kuliah Statistika Bisnis dan banyak melakukan penelitian di bidang Pemasaran.

## BIODATA PENULIS

**Derry Nugraha., M.Pd.**
Dosen Matematika dan Statistika
Universitas Linggabuana PGRI Sukabumi

Derry Nugraha, M.Pd., lahir di Sukabumi, 14 Januari 1990. Dari ayah bernama Duduh Abdullah dan Ibu bernama Amelia. Ia memiliki seorang istri bernama Desmianty Purwanty, S.Pd. Penulis bertempat tinggal di Kelurahan Baros Kecamatan Baros Kota Sukabumi Provinsi Jawa Barat. Telah menyelesaikan studi strata satu di Program Studi Pendidikan Ekonomi (2008 – 2012).

Lulus strata dua di Program Studi Magister Pendidkan Matematika Institut Keguruan dan Ilmu Pendidikan Siliwangi Bandung (2016-2018). Sedang menyelesaikan studi program doktoral pada Universitas Pendidikan Indonesia Bandung Program Studi Pascasarjana Doktor Pendidikan Matematika.

Karirnya dimulai sebagai Dosen Tetap Yayasan di Universitas Linggabuana PGRI Sukabumi (2016 – sekarang). Dosen tidak tetap di beberapa Perguruan Tinggi di Kota Sukabumi. penulis menjabat salah satu Editor Jurnal Nasional pada Journal of Education and Culture sejak tahun 2022,

Editor pada Jurnal Multidisiplin Ilmu Indragiri sejak tahun 2023, Editor Jurnal Ikhlas : Pengabdian Pada Masyarakat sejak tahun 2023, dan Jurnal Pengabdian Inovasi Dosen dan Mahasiswa sejak tahun 2023. Penulis juga aktif dalam beberapa penulisan artikel sejak tahun 2016 dan menjadi narasumber dalam beberapa pertemuan ilmiah tingkat daerah dan nasional.

## BIODATA PENULIS

**Tedi Hartoyo, Ir. MSc.**

Dosen Program Studi Agribisnis Fakultas Pertanian Universitas Siliwangi. Dosen Statistika Jurusan Kimia Analis BTH Tasikmalaya. Dosen Statistika Keperawatan Respati Tasikmalaya. Dosen Statistika keperawatan Gigi Tasikmalaya.

Penulis lahir di Tasikmalaya tanggal 26 Juli 1962. Penulis adalah Dosen Program Studi Agribisnis Fakultas Pertanian Universitas Siliwangi. Menyelesaikan pendidikan S1 pada Jurusan Statistika IPB Bogor dan melanjutkan S2 pada Agriculture Development, Gent Unversity Belgium. Saat ini penulis mengampu beberapa mata kuliah, diantaranya Matematika, Statistika, Ekonometrika, Metode Penelitian Sosial Ekonomi dan Perancangan Percobaan.

# BIODATA PENULIS

**Unang Atmaja, Ir. MSc.**
Dosen Program Studi Agribisnis Fakultas Pertanian Universitas Siliwangi

Penulis lahir di Tasikamalaya tanggal 17 Agustus 1964. Penulis adalah Dosen Program Studi Agribisnis Fakultas Pertanian Universitas Siliwangi. Menyelesaikan pendidikan S1 pada Jurusan Perikanan Universitas Padjadjaran dan melanjutkan S2 pada Agriculture Development, Gent Unversity Belgium. Saat ini penulis mengampu beberapa mata kuliah, diantaranya Manajemen Agribisnis,Ekonomi Sumberdaya Alam dan Lingkungan.

## BIODATA PENULIS

**Dodi, S.Pd., M.Pd.**
Dosen Program Studi Ilmu Kelautan
Fakultas Ilmu Pengetahuan Alam dan Kelautan
Universitas OSO

Penulis berasal dari Padang Tikar dan lahir pada tanggal 20 Agustus 1990. Penulis merupakan seorang dosen tetap di Program Studi Ilmu Kelautan, Fakultas IPA dan Kelautan, Universitas OSO. Ia menyelesaikan pendidikan S1 di program studi Pendidikan Matematika di UNTAN sebelum melanjutkan ke jenjang S2 di Program Studi Pendidikan Matematika di UNTAN. Penulis menekuni dunia kepenulisan pada bidang Matematika dan Sains, serta bidang Pendidikan.

Penulis berharap, karya yang diterbitkannya dapat memberikan inspirasi serta manfaat bagi orang lain untuk belajar matematika dan sains. Kontak penulis tercantum sebagai berikut; email: dodi_a1@yahoo.com / dodi@oso.ac.id; No Hp/ Wa: 089694156039.

## BIODATA PENULIS

**Fitriana Dina Rizkina, S.T.P., M.Sc.**
Dosen Program Studi Teknologi Industri Pertanian
Fakultas Pertanian
Universitas Muhammadiyah Jember

Penulis merupakan akademisi dan praktisi di bidang Manajemen Industri Pertanian (Agroindustri) yang lahir di Jember tanggal 7 Maret 1994. Penulis adalah dosen tetap pada Program Studi Teknologi Industri Pertanian Fakultas Pertanian, Universitas Muhammadiyah Jember. Menyelesaikan pendidikan S1 pada Program Studi Teknologi Industri Pertanian, Institut Pertanian Bogor dan melanjutkan S2 pada Program Studi Teknologi Industri Pertanian, Universitas Gadjah Mada. Saat ini penulis juga merupakan mahasiswa Doktoral (S3) di Ehime University, Jepang. Penulis menekuni bidang Manajemen Industri Pertanian (Agroindustri): Manajemen Rantai Pasok, Manajemen Risiko, Manajemen Bisnis, Inovasi dan Sosial Ekonomi. Penulis dapat dihubungi melalui email: fitrianadina@unmuhjember.ac.id.