STATISTIKA

Untuk Ekonomi dan Bisnis

☑ **Ir. S. Benny Pasaribu, M.Ec., Ph.D.**
Fakutas Ekonomi dan Bisnis
Universitas Trilogi
Ph.D. dalam Ilmu Ekonomi Industri
Ottawa University, Kanada

☑ **Rizqon Halal Syah Aji, M.Si., Ph.D.**
Fakutas Ekonomi dan Bisnis
Universitas Islam Negeri Syarif Hidayatullah Jakarta
Ph.D. dalam Ilmu Ekonomi Pembangunan
Universitas Kebangsaan Malaysia (UKM)

☑ **Dr. Kabul Wahyu Utomo, M.Si.**
Fakultas Ekonomi dan Bisnis
Universitas Trilogi
Doktor dalam Ilmu Manajemen
Universitas Gajah Mada (UGM)

☑ **Dr. Aty Herawati**
Fakultas Ekonomi dan Bisnis
Universitas Trilogi
Doktor Manajemen Bisnis Bidang Keuangan
IPB University

EDU PUSTAKA

METODOLOGI PENELITIAN
Untuk Ekonomi dan Bisnis

Penulis : Ir. S. Benny Pasaribu, M.Ec., Ph.D
Rizqon Halal Syah Aji, M.Si., Ph.D
Dr. Kabul Wahyu Utomo, M.Si
Dr. Aty Herawati

Editor : ...

Diterbitkan oleh:
Penerbit **EDU PUSTAKA**
Anggota IKAPI

ISBN:
Hal. xii + 320, Uk. 15,5 x 23 cm
Cetakan Pertama, 2021

Pemasaran:
Jl. Haji Karim No. 70 Setu, Cipayung, Jakarta Timur 13880
Telefaks. (021-70300534)
Email: penerbitedupustaka@gmail.com

Kata Pengantar

Penelitian ekonomi dan bisnis yang pendekatannya kuantitatif memerlukan alat analisis berupa ilmu statistika. Ilmu statistika merupakan ilmu yang menampilkan fakta-fakta normatif (naturalistik), empirikal (sesuai kondisi data), dan praktis (dapat di implemaentasikan secara praktikal). Tiga hal tersebut divisualisasi melalui tabel, diagram ataupun grafik yang di dalamnya memiliki makna tertentu dalam setiap analisisnya. Ilmu statistika juga bermanfaat dalam menyusun model perumusan hipotesis, pengembangan model penelitian maupun menyusun disain penelitian dalam penentuan sampel dan analisis data.

Buku ini, hadir dalam mengisi khasanah pengembangan ilmu statistik untuk praktik penelitian pada lingkup ekonomi dan bisnis. Keunggulan buku ini adalah di bahas dengan bahasa sederhana serta melingkupi topik-topik bahasan ilmu statistik yang paling sederhana dari deskripstif statistik hingga statistik inferensia.

Buku ini disusun atas dasar kebutuhan pembelajaran dan praktik-praktik penelitian yang dialami oleh para penulis, yang setiap harinya menekuni sebagai Dosen, sekaligus peneliti dan praktisi. Oleh karena itu buku ini mencoba mengeksplorasi pengalaman praktis yang kemudian dirumuskan menjadi sebuah buku agar dapat merangkumi penjelasan teoritik untuk memayungi hal-hal praktik penelitian yang biasa dilakukan di kampus sekaligus dalam dunia kerja.

Besar harapan kami, buku ini dapat melengkapi khasanah dan sekaligus membantu mahasiswa dalam memahami ilmu statistik pada lingkup penelitian ekonomi dan bisnis.

Tim Penulis

Daftar Isi

Bab 1
Pendahuluan

1.1 PENGANTAR

Bidang ilmu statistik mempunyai dua jenis pendekatan dalam melakukan analisis yaitu statistik deskriptif dan statistik inferensia. Fokus utama bab ini adalah mengenalkan kepada pembaca apa itu statistik melalui definisi atau pengertian secara umum, kegunaan dan fungsi dari setiap metode statistik, perbedaan antara unsur-unsur dalam statistik yaitu data, populasi, sampel dan peubah, dan apa itu statistik parametrik dan nonparametrik. Sehingga, diakhir proses pembelajaran dari bab ini mahasiswa mengetahui konsep dasar ilmu statistik serta dapat menyajikan data dalam bentuk tabel, grafik atau diagram.

1.2 PENGERTIAN STATISTIK

Secara definisi statistik dapat diartikan sebagai kumpulan angka atau informasi yang dapat dijelaskan kembali secara baik melalui interpretasi dan pemahaman setelah melalui proses pengolahan angka dan informasi dengan tujuan supaya dapat lebih mudah dipahami oleh pembaca. Namun, dalam perkembangannya statistik menjadi lebih dikenal dengan pengolahan informasi yang berbentuk angka sehingga dapat dianalisis dengan menggunakan pendekatan parameter tertentu supaya menghasilkan gambaran situasi, kejadian atau peristiwa. Proses analisis angka-angka tersebut biasa dikenal dengan pengolahan data kuantitatif. Umumnya, kelompok data kuantitatif terdiri dari data kependudukan, kenaikan dan penurunan harga dalam perdagangan, bidang pendidikan seperti jumlah pelajar, jumlah tenaga kerja dan hasil pertanian. Secara spesifik, data kependudukan meliputi jumlah penduduk baik jumlah laki-laki maupun perempuan, tingkat kelahiran dan kematian, rata-rata umur penduduk, dan angka penduduk miskin. Statistik dalam perdagangan meliputi tren kenaikan harga barang di pasar, kemampuan daya beli masyarakat serta fluktuatifnya inflasi sebuah negara. Statistik untuk bidang pendidikan pula terdiri dari jumlah bangunan sekolah, jumlah guru, jumlah sekolah, jumlah lulusan dan tenaga pengajar dan lain-lain.

Sekumpulan informasi seperti yang dijelaskan di atas disebut dengan Data dalam statistik. Dalam arti sempit data-data tersebut merupakan data statistik. Sedangkan dalam arti luas statistik adalah bidang ilmu yang mempelajari tentang pengumpulan dan penyajian data, proses analisis data, membuat interpretasi dan memberikan kesimpulan dan saran dari hasil analisis tersebut. Sehingga, data yang diolah dapat memberikan informasi baru yang relevan dengan peristiwa yang sedang diamati.

1.3 STATISTIK DESKRIPTIF DAN INFERENSIA

Sebagaimana yang sudah disebutkan di awal, statistik terdiri dari dua jenis yaitu statistik deskriptif dan statistik inferensia. Statistik deskriptif (statistik deduktif) secara definisi merupakan proses analisis melalui penggambar atau pendeskripsian kejadian berdasarkan sifat atau kriteria dari sekelompok data yang diamati. Dalam proses analisisnya, kesimpulan yang diberikan bukanlah kesimpulan secara general karena kelompok data yang diuji adalah sebagian atau beberapa sampel dari seluruh populasinya. Proses pengerjaan analisis deskriptif mencakupi pengumpulan data, penyusunan dan pengolahan serta penyajian data dan kemudian dianalisis supaya menggambar suatu kejadian dan dapat dijelaskan secara ringkas, padat dan jelas melalui penarikan kesimpulan. Sedangkan, statistik inferensia (statistik induktif) mempunyai aturan yang harus dipenuhi sebelum data yang diperoleh dapat dianalisis. Kegunaan dari analisis menggunakan pendekatan statistik inferensia adalah dapat memberikan ramalan (*forecasting*) dari hasil proses analisis untuk pengambilan keputusan di masa depan, membuat taksiran untuk suatu kejadian. Data yang digunakan merupakan data secara keseluruhan yang dipilih berdasarkan peubah yang ingin dianalisis. Biasanya, kelompok data ini diambil dari sekumpulan data yang juga menjadi subjek kajian (populasi). Manfaat penggunaan statistik inferensia adalah dapat memberikan ramalan melalui pendugaan suatu populasi yang kemudian diuji dengan menggunakan hipotesis dengan parameter yang telah ditetapkan. Sehingga, teori tentang peluang suatu kejadi masih erat kaitannya dengan statistik inferensia. Secara proses pengerjaan, statistik inferensia adalah proses keberlanjutan dari statistik deskriptif. Oleh sebab itu, sebelum lebih lanjut memahami statistik inferensia ada baiknya mempelajari dan memperdalam konsep statistik deskriptif.

1.4 FUNGSI DAN KEGUNAAN STATISTIK

Fungsi dan kegunaan statistik dalam dunia ekonomi dan bisnis adalah sebagai alat bantu. Masyarakat dan pemerintah sebagai pelaku ekonomi harus dapat memahami konsep statistik terutama dalam hal pengambilan keputusan. Sehingga, dengan menggunakan alat bantu dalam proses analisis data akan memudahkan pelaku ekonomi untuk pengumpulan, pengolahan dan menganalisis dan menarik sebuah kesimpulan dari fenomena yang diamati. Berdasarkan hasil pengolahan data tersebut, pelaku ekonomi memperoleh hal-hal berikut:

1. Membuat keputusan dari suatu kejadian atau peristiwa ekonomi secara umum atau secara khusus.
2. Mengetahui perkembangan terkini dari fenomena yang sedang diteliti dari waktu ke waktu.
3. Menyusun laporan, baik keuangan maupun manajerial, baik secara deskriptif maupun kuantitatid secara baik, jelas dan teratur.
4. Mengetahui dampak suatu kegiatan sosial, budaya atau politik terhadap perekonomian.
5. Dapat membuat ramalan terhadap kemungkinan dampak dari hubungan bidang-bidang sektor lain terhadap kegiatan ekonomi dan bisnis yang dapat dipertanggung jawabkan hasil temuannya secara ilmiah.

1.5 DATA, POPULASI DAN SAMPEL

Data merupakan sekumpulan informasi baik berupa angka atau tidak berbentuk angka dari suatu fenomena atau kejadian. Data yang diperoleh disebut juga dengan **data statistik.** Data statistik tersebut juga terdiri dari dua jenis. Bagi setiap data yang berbentuk angka dan mempunyai nilai atau setelah dilakukan kuantifikasi berupa pemberian kode atau label disebut dengan data kuantitatif. Bagi setiap data informasi berupa keterangan, penjelasan dari sebuah kejadian yang tidak berbentuk angka dikenal dengan data kualitatif. Data kualitatif merupakan sekumpulan data yang tidak dapat diukur menggunakan skala numerik. Tetapi data ini dapat dikelompokkan menjadi satu kategori dengan klasifikasi kalimat penjelas yang sama. Contoh data kualitatif seperti jenis kelamin, warna kesukaan dan brand yang diminati. Sedangkan, untuk data kuantitatif merupakan data yang mempunyai nilai atau data kualitatif yang telah melalui proses pengkodeaan nilai yang pada akhirnya mempunyai nilai skala tersendiri. Contohnya adalah pendapat dari responden diambil menggunakan 3 skala, setuju, netral dan tidak setuju. Apabila dinumerikan menjadi 1 untuk setuju, 2 netral dan 3 tidak setuju. Contoh lainnya adalah jumlah mahasiswa ekonomi dan bisnis yang hadir tepat waktu, pengendara sepeda motor yang masuk ke lingkungan kampus dan lain sebagainya. Data kuantitatif juga terdiri dari dua jenis yaitu data diskrit dan data kontinu. Data diskrit adalah hasil pencacahan dan data kontinu adalah hasil dari suatu pengukuran.

Populasi adalah sekumpulan unit atau individu dalam penelitian yang dapat dibedakan berdasarkan jenis objek penelitian. Unit tersebut mempunyai karakteristik pembeda dan inging diketahui sebagai objek penelitian. Objek penelitian yang dimaksud seperti institusi atau lembaga tertentu, kelompok masyarakat atau orang, produk baik barang atau jasa, berat, besaran atau tinggi dan lain-lain. Sedangkan karakteristik dari obejk penelitian berupa sifat, ciri atau fenomena berulang yang terjadi sebagai sebuah pengamatan. Informasi berupa keterangan dari suatu kejadian kemudian dikumpulkan dan menghasilkan data statistik. Peneliti harus menentukan populasi yang akan diteliti. Selama proses penelitian, batasan terhadap populasi harus dapat dijelaskan sehingga tidak menimbulkan bias bagi pembaca. Ukuran populasi juga perlu diperhatikan karena mempertimbangkan luas daerah penelitian yang mana semakin luas daerah penelitian maka semakin besar populasi yang akan diobservasi.

Unsur terakhir yang harus ditentukan dalam sebuah penelitian adalah besaran **Sampel** penelitian**.** Sampel merupakan bagian kecil dari seluruh populasi. Biasanya yang akan diselidik adalah ciri khas atau karakteristik dari setiap sampel.

1.6 SAMPEL STATISTIK DAN PARAMETER POPULASI

Statisik dan parameter merupakan dua istilah yang saling terkait terutama dalam proses pengujian hipotesis dan pendugaan parameter populasi. Statistik adalah karakteristik dari sampel-sampel penelitian yang diukur menggunakan perhitungan tertentu dan menghasilkan nilai rata-rata sampel ($\bar{x}$), simpangan baku (s), koefisien korelasi (r), koefisien regresi (b), proporsi sampel ($\hat{p}$), beda dua rata-rata sampel (X_1-X_2), dan beda dua proporsi sampel ($\hat{p}_1 - \hat{p}_2$).

Sedangkan, parameter merupakan suatu ukuran dari populasi berdasarkan sifat atau karakteristiknya seperti rata-rata populasi (μ), simpangan baku populasi (σ), koefisien korelasi populasi (σ), koefisien model regresi populasi (β), proporsi populasi (P), beda dua rata-rata populasi ($\mu_1 - \mu_1$), dan beda dua proporsi populasi ($P_1 – P_2$).

1.7 PEUBAH

Peubah mempunyai besaran nilai yang bervariasi dan keberagaman ciri atau sifat. Peubah juga terdiri dari dua jenis, peubah kualitatif dan peubah kuantitatif. Pada peubah kualitatif lebih memperhatikan sifat atau ciri yang dimiliki sehingga tidak perlu menggunakan angka bilangan.

Contoh sederhana peubah kualitatif adalah (1) jumlah pemilih berdasarkan jenis kelamin, dimana pemilih laki-laki berjumlah 300 orang (60%) dan pemilih perempuan berjumlah 200 orang (40%); (2) dari 3 jenis mangga yang dipasarkan, mangga madu, mangga gincu dan mangga muda, yang paling diminati oleh pembeli adalah mangga madu yang terjual 10 kg (50%) sedangkan mangga gincu dan mangga madu terjual dengan jumlah yang sama masing-masing 5 kg (25% untuk mangga gincu dan 25% untuk mangga madu).

Sebaliknya, peubah kuantitatif adalah sekelompok data baik bilangan atau numerik dengan variasi nilai yang dapat diolah menghasilkan nilai frekuensi. Sama halnya dengan data kuantitatif, peubah kuantitatif terdiri dari dua jenis peubah, yaitu peubah diskrit dan kontinu. Peubah diskrit selalunya berbentuk bilangan cacah dan mempunyai perhitungan yang lebih sederhana. Contoh peubah diskrit adalah (1) jumlah mahasiswa fakultas ekonomi dan bisnis tahun ajaran 2021 sebanyak 800 orang, dimana mahasiswa dengan jurusan ilmu ekonomi pembangunan berjumlah 400 mahasiswa, ilmu akuntasi berjumlah 200 mahasiswa dan ilmu manajemen sebanyak 200 mahasiswa; (2) penjualan mobil berdasarkan brand di showroom mobil PT. XYZ selama 3 bulan adalah sebagai berikut, Mobil dengan brand Honda berhasil terjual 10 unit setiap bulan, mobil dengan brand BMW berhasil terjual 5 unit bulan pertama, 7 unit bulan kedua dan 4 unit bulan ketiga. Peubah kontinu diperoleh dari hasil pengukuran sehingga datanya dinyatakan dalam bentuk bilangan desimal (pecahan). Contoh peubah kontinu adalah berat badan.

Bab 2
Pengumpulan Data

2.1 PENGANTAR

Pada bab ini akan dijelaskan tentang apa itu pengumpulan data dan metode-metode dasar dalam ilmu statistik. Pengumpulan data merupakan langkah awal ketika memulai proses penelitian. Hal ini sangat penting dilakukan sebelum memasuki proses analisis. Sehingga, berdasarkan data yang berhasil dikumpulkan kita dapat menentukan jenis analisis yang sesuai supaya akurasi hasil penelitian semakin baik dan tepat.

2.2 TAHAPAN PENELITIAN DALAM METODE STATISTIK

Setiap peneliti sebaiknya mengikut tahapan yang sesuai selama menjalankan sebuah penelitian. Ini akan memudahkan pekerjaan terutama untuk pengambilan kesimpulan. Oleh sebab itu tahapan yang harus dilakukan sebagai berikut:

1. Pengumpulan data (Data collection)
2. Penyusunan data (Data organization)

3. Pengolahan data (Data tabulation)
4. Analisis data (Data analysis)
5. Interpretasi data (Data interpretation)

2.2.1 Pengumpulan Data

Proses awal dalam penelitian adalah melakukan pengumpulan data. Pengumpulan data dapat dilakukan dengan dua cara yaitu mengambil sampel, baik dari populasi atau sampel data dan melakukan sensus terhadap sampel penelitian. Pengambilan data sampel atau *sampling* adalah cara yang paling mudah dilakukan. Data yang dikumpulkan hanya sebagian dari seluruh populasi objek penelitian. Data yang akan diperoleh adalah data perkiraan atau nilai estimasi, kemudian ditaksir ke dalam masing-masing kelompok data berdasarkan sifat atau karakteristik dari populasi tersebut. Sampel yang baik harus dapat mewakili (representatif) populasi penelitian sehingga menghasilkan perkiraan yang tepat. Penentuan jumlah sampel dapat dilakukan dengan menggunakan beberapa metode, salah satunya adalah perhitungan menggunakan rumus yang akan dijelaskan dalam bab sampling.

Cara bengumpulan data kedua adalah melalui sensus. Badan Pusat Statistik Republik Indonesia selalu menggunakan cara ini untuk setiap kajian penelitian yang sedang dilakukan. Hal ini karena cara sensur dinilai lebih akurat karena proses pengumpulan dilakukan secara langsung melalui pencatatan informasi data dari seluruh anggota objek penelitian. Sehingga, seluruh anggota objek penelitian ini juga disebut sebagai kelompok populasi. Dikarenakan semua anggota objek penelitian harus diteliti satu per satu, sehingga bagi populasi dalam jumlah besar, pengumpulan data melalui cara sensus membutuhkan waktu yang panjang dan pengumpul data (enumerator) yang tidak sedikit. Walaupun biaya, waktu dan tenaga yang diperlukan sangat besar dalam proses pengumpulan data ini, namun data yang diperoleh merupakan data yang riil dan benar (true value).

2.2.2 Penyusunan Data

Setelah data terkumpul, langkah selanjutnya yang harus dilakukan adalah menyusun data dan mengelompokkan masing-masing data berdasarkan kategori. Hal ini bertujuan supaya data mudah dilihat, dibaca dan difahami. Penyusunan data dapat dilakukan melalui 3 tahapan, yaitu

1. Mengedit, artinya setiap data kembali diperiksa dengan menyesuaikan pertanyaan dengan jawaban atau mengelompokkan data berdasarkan tahun yang sama. Tujuannya supaya data tersebut telah diisi dengan benar (pada data kuesioner) dan tidak terjadi kesalahan penempatan dalam tahun (pada data perusahaan).
2. Mengklasifikasi, artinya peneliti harus memisahkan setipa butih2 data yang telah diedit berdasarkan kelompok peubahnya masing-masing.

2.2.3 Pengolahan Data

Dalam poin 2 di atas, setelah pengelompokkan data telah dilakukan kita dapat mengolah data dengan menggunakan metode sederhan seperti menghitung nilai rata-rata, media dan modus. Pengolahan data juga termasuk dalam mengelompokkan data. Perbedaannya adalah pengelompokkan data pada pengolahan data adalah menginput data kembali sesuai dengan kolom dan baris. Sehingga, ketika data akan diumumkan pembaca lebih mudah melihat data secara visual dengan merujuk pada tabel atau kolom dan pada akhirnya dapat menghasilkan penyajian dalam bentuk grafik atau diagram.

2.2.4 Analisis Data

Data yang telah dikelompokkan dan diolah dapat digunakan dan diinput ke dalam alat analisis. Terdapat pelbagai alat analisis yang dapat digunakan seperti SPSS, eViews dan Stata. Dari data yang terkumpul, disusun dan kemudian dianalisis maka peneliti memperoleh gambaran menyeluruh terhadap fenomena berdasarkan data yang dianalisis.

2.2.5 Interpretasi Data

Keseluruhan informasi baik dari pengolahan maupun analisis data dapat diinterpretasi supaya menghasilkan sebuah kesimpulan dan saran kepada pihak yang terkait dengan objek penelitian. Sekiranya berhubungan dengan pemerintah maka pemerintah dapat menggunakan hasil penilitian sebagai bahan dasar pertimbangan pembuatan kebijakan.

2.3 SUMBER DATA

Berdasarkan dari cara perolehannya, sebuah data dikelaskan ke dalam 2 kelas, yaitu kelas data primer dan sekunder. Data primer merupakan data yang belum diolah atau dipergunakan untuk keperluan lain dan bersumber langsung objek penelitian. Contoh sederhana data primer adalah laporan keuangan firma sebelum dilakukannya pengauditan. Selain itu, peneliti yang sedang melakukan pengumpulan data mahasiswa di kampus selama 1 bulan pengamatan juga merupakan sumber data primer. Pengamatan dilakukan dengan cara observasi dan wawancara setiap mahasiswa secara langsung. Contoh lainnya adalah penelitian terhadap kepuasan pengguna kartu sim pada telefon pintar. Sebelum dilakukannya penelitian, maka perlu dilakukan pengumpulan data terkait tingkat kepuasan pelanggan terhadap harga, jaringan dan masa berlaku kartu yang digunakan. Sehingga, dari ketetapan ini, peneliti sudah menetapkan jumlah sampe responden yang akan diwawancara dengan menggunakan kuesioner.

Data sekunder diperoleh secara langsung dari objek penelitian dalam hal ini seperti publikasi resmi yang dikeluarkan oleh objek penelitian. Data yang dikeluarkan merupakan data yang telah diolah. Sehingga, kita bisa menggunakan data tersebut tanpa mempertimbangkan aspek lain dalam perhitungan sampel. Contohnya adalah jumlah penduduk Indonesia. Data yang dikeluarkan oleh Badan Pusat Statistik melaporkan bahwa pada tahun 2-1- penduduk Indonesia berjumlah 237.6 juta jiwa dan angka ini terus mengalami peningkatan. Hal ini dapat dilihat pada tahun 2021, 11 tahun kemudian, jumlah penduduk Indonesia menjadi sebanyak 270.6 juta jiwa.

2.4 KELOMPOK DATA

Dilihat dari pengelompokkan data, data dibedakan ke dalam dua kelompok yaitu data penampang silang (*cross* section) dan data berkala (*time series*).

Data penampang silang (*cross section data*) adalah daya dari objek penelitian yang menggambarkan karakteristik data penelitian seperti Tingkat fertilitas penduduk di Indonesia, perubahan harga bahan sembako dengan menggunakan harga dasar dan indeks pembangunan manusia Indonesia periode tahun 2016-2020.

Data berkala (*time series data*) biasanya dikenal dengan data runtut waktu yang mana data ini dikelompokkkan mengikut waktu ke waktu selama periode tahun penelitian. Data tersebut dapat menjelaskan gambaran perkembangan suatu peristawa dari objek peneltiain. Sebagai contoh, (1) perkembangan nilai impor dan ekspor di Indonesia selama 10 tahun berawal dari 2011-2021.

2.5 UKURAN DATA

Berdasarkan skala pengukurannya, data statistik dapat dibedakan menjadi empat macam yaitu: data nominal, data ordinal, data interval dan data rasio.

2.5.1 Data Nominal

Data nominal disusun berdasarkan kategori seperti kedudukan data. Dalam tabel data nominal tujuan diberikan kode pada data untuk memperjelas pada masing-masing kelompok data. Contoh data nominal adalah dari seluruh mahasiswa ilmu ekonomi fakultas ekonomi dan bisnis yang berjumlah 400 mahasiswa, jumlah laki-laki sebanyak 212 mahasiswa dan perempuan sebanyak 188 mahasiswa. Dari jumlah tersebut setelah dilakukan pengumpulan data, mahasiswa yang bersetuju terhadap rencana pembelajaran secara daring berjumlah 70 persen sedangkan 30 persen lainnya memilih tidak setuju.

2.5.2 Data Ordinal

Data ordinal dapat disusun dengan mempertimbangkan urutan atau kedudukan dari sekelompok data. Kedudukan dalam data dapat ditulis sebagai peringkat. Peringkat tersebut nantinya akan digunakan untuk mengurutkan angka sesuai besar kecilnya sebuah nilai. Contoh untuk data ordinal seperti berikut, dari data mahasiswa jurusan ilmu ekonomi di fakultas ekonomi dan bisnis, berdasarkan kuesioner yang telah dikumpulkan, setiap jawaban pada masing-masing pertanyaan diberi kode 1 untuk penilaian setuju, kode 2 untuk netral dan kode 3 untuk tidak setuju. Pertanyaan yang dapat diajukan seperti berikut: "*Apakah menurut anda pelayanan yang telah diberikan oleh pegawai fakultas sangat membantu*?" Dari hasil pengumpulan data diperoleh respons yang beragama dan berdasarkan pengelompokkan data diperoleh data sebagai berikut, mahasiswa yang memberi penilaian setuju berjumlah 319 mahasiswa, netral sebanyak 41 mahasiswa dan tidak setuju berjumlah 40 mahasiswa.

2.5.3 Data Interval

Data interval pada data statistik disusun dengan berdasarkan kepada kategori data. Data interval juga mempunyai sifat yang sama dengan data nominal dan data ordinal, namun perbedaan pada data interval adalah jarak atau frekuensi yang sama untuk masing-masing kategori data. Sehingga, dalam proses penjarakkan nilai dari responden operasi aritmatika berlaku pada data ini. Contoh pada borang kuesioner yang dibagikan kepada 400 mahasiswa ilmu ekonomi fakutlas ekonomi dan bisnis. Salah satu pertanyaan yang diberikan berisi tentang frekuensi sesorang tidak mengikuti kelas pada satu mata kuliah tertentu. Pertanyaan seperti berikut "Berapa kali anda tidak hadir ke dalam kelas statistik?". Tentunya jawaban dari mahasiswa sangat beragam karena mempertimbangkan syarat bagi setiap mahasiswa untuk tidak mengikuti perkuliahan adalah sebanyak 4 kali tetapi tidak menutup kemungkinan ada yang menjawab lebih dari 4 kali. Respons yang diperoleh dari para mahasiswa ilmu ekonomi adalah sebagai berikut: yang menjawab tidak pernah sebanyak 289 mahasiswa, yang menjawab bolos 2 kali sebanyak 111 mahasiswa, tidak hadir sebanyak 3 kali sebanyak 30 mahasiswa, 40 mahasiswa yang tidak datang ke kelas sebanyak 4 kali, tidak datang ke kelas sebanyak 5 kali berjumlah 20 mahasiswa, dan 10 mahasiswa menjawab tidak menghadiri perkuliahan sebanyak 6 kali.

2.5.4 Data Rasio

Data rasio diperoleh dengan membuat perbandingan antara nilai dasar dan nilai pembanding. Sifat pada data rasio meliputi sifat pada tiga data lainnya yaitu nominal, ordinal dan interval. Namun, yang membedakan data rasio dengan data lainnya adalah nilai nol pada data rasio mempunyai nilai dan biasanya dianggap ketiadaan atau tidak berkaitan satu sama lain. Contoh sederhan untuk data rasio adalah seperti berikut, Clara ingin membeli baju di toko baju dekat rumah dengan harga Rp50.000, kemudian ketika berjalan di toko baju di sebuah mal, Clara melihat baju dengan gaya yang sama seharga Rp200.000. Dari informasi ini kita dapat simpulkan bahwa harga baju di toko dekat rumah empat kali lebih murah dari harga baju di mal.

Bab 3
Pengolahan Data

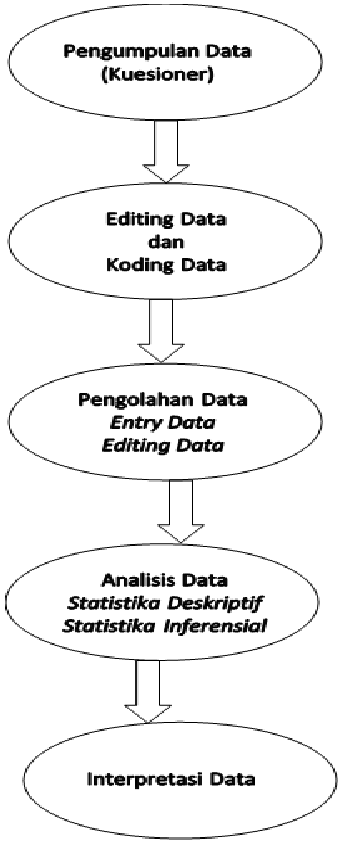
3.1 PENGANTAR

Dalam bab ini akan diuraikan tentang prosedur pengolahan data, tahapan pengolahan data dan contoh syarat pengujian dalam pengolahan data penelitian. Pengolahan data itu sendiri bertujuan untuk mempersiapkan data sehingga memudahkan analisa data. Oleh karena itu, pengolahan serta penyesuaian data harus dilakukan sebelum data tersebut dapat digunakan bagi tujuan analisis.

3.2 PROSEDUR PENGOLAHAN DAN ANALISIS DATA

Tujuan pokok dilaksanakannya penelitian adalah untuk *menjawab* pertanyaan- pertanyaan penelitian. Untuk mencapai tujuan pokok tersebut antara lain harus melalui proses pengolahan dan analisis data. Alur kerjanya, yang dimulai dari pengumpulan hingga interpretasi data dapat dilihat pada ilustrasi berikut ini:

GAMBAR 3.1 *Ilustrasi Proses Pengolahan dan Analisis Data*



3.2.1 Pengumpulan dan Pengolahan Data

Pengumpulan Data

Sebelum melakukan pengolahan data, ada beberapa tahap yang harus dilakukan. Sedangkan setelah analisis data yaitu suatu proses penyederhanaan data, maka dapat dilakukan interpretasi data dengan mudah. Kuesioner merupakan alat pengumpul data yang digunakan untuk survai, guna memudahkan proses selanjutnya, sebaiknya dalam kuesioner telah tersedia kolom untuk koding.

Editing Data

Data lapangan yang ada dalam kuesioner perlu diedit, tujuan dilakukannya editing adalah untuk: (1) Melihat lengkap tidaknya pengisian kuesioner. (2) Melihat logis tidaknya jawaban. (3) Melihat konsistensi antar pertanyaan.

Koding Data

Dilakukan untuk pertanyaan-pertanyaan: (1) Tertutup, bisa dilakukan pengkodean sebelum ke lapangan. (2) Setengah terbuka, pengkodean sebelum dan setelah dari lapangan. (3) Terbuka, pengkodean sepenuhnya dilakukan setelah selesai dari lapangan.

Pengolahan Data

Paling tidak ada dua hal yang perlu dilakukan ketika melakukan pengolahan data: (1) *Entry data*, atau memasukan data dalam proses tabulasi. (2) Melakukan editing ulang terhadap data yang telah ditabulasi untuk mencegah terjadinya kekeliruan memasukan data, atau kesalahan penempatan dalam kolom maupun baris tabel.

3.2.2 Analisis dan Interpretasi Data

Hal penting yang perlu diingat dalam melakukan analisis data adalah mengetahui dengan tepat penggunaan alat analisis, sebab jika kita tidak memenuhi prinsip-prinsip dari pemakaian alat analisis, walaupun alat analisisnya sangat canggih, hasilnya akan salah diinterpretasikan dan menjadi tidak bermanfaat untuk mengambil suatu kesimpulan. Model-model statistika untuk keperluan analisis data telah begitu berkembang, dari model-model statistika deskriptif hingga ke statistika inferensial non parametrik dengan persyaratan yang lebih "lunak " dibandingkan dengan statistika parametrik yang sangat ketat dengan persyaratan-persyaratan tertentu dan sulit dipenuhi dalam kerangka penelitian sosial.

Ketika kita memutuskan untuk melakukan analisis data menggunakan alat statistika, ada beberapa hal yang perlu diperhatikan antara lain:

1. Dari mana data diperoleh, apakah berasal dari sampel (melalui proses sampling) atau dari populasi (dengan cara sensus)

2. Jika berasal dari sampel apa teknik sampling yang digunakan, apakah termasuk kelompok sampling probabilitas atau non probabilitas.
3. Memakai skala apa data diukur, apakah menggunakan skala nominal, ordinal, interval, atau rasio.
4. Bagaimana hipotesis yang dibuat apakah perlu dilakukan pengujian satu arah atau dua arah kalau memakai statistika inferensial.

3.2.3 Statistika Deskriptif sebagai Alat Analisis

Statistika Deskriptif merupakan metode atau alat analisis yang biasa digunakan untuk menyederhanakan data agar mudah dipahami. Penyajiannya bisa dalam bentuk tabel, baik tabel frekuensi maupun tabel silang atau dalam bentuk diagram dan grafik seperti diagram batang, kurva dll.

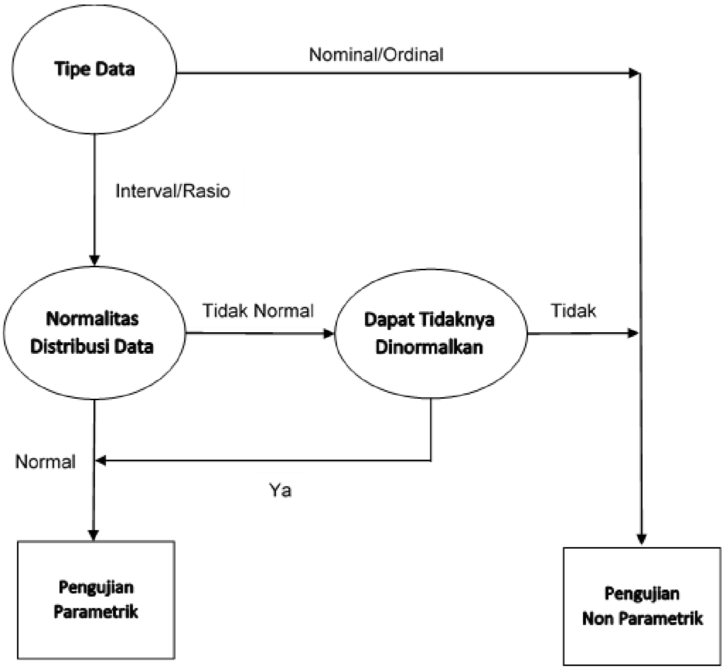
Statistika deskriptif dapat diterapkan baik untuk data yang berasal dari sampel maupun populasi, juga untuk sampel yang diambil dengan sampling probabilitas maupun non probabilitas, serta bisa digunakan untuk semua skala pengukuran dari mulai yang paling lemah (nominal) hingga skala rasio.

Statistika Deskriptif sering digunakan untuk mengukur *gejala pemusatan*, dan *dispersi atau simpangan data*. Termasuk ukuran gejala pusat antara lain:: modus, median, persentil, mean atau rata-rata. Tergolong ukuran dispersi data antara lain: rentang (maksimum - minimum), deviasi standard, koefisien variasi. Jika dikaitkan dengan skala pengukuran dari data yang dianalisis, statistika deskriptif yang cocok digunakan adalah:

1. Skala Nominal: Modus, Frekuensi
2. Skala Ordinal: Median, Persentil, Rentang
3. Skala Interval: Mean, Deviasi Standard
4. Skala Rasio: Mean, Koefisien Variasi (ukuran dispersi relatif)

3.2.4 Statistika Nonparametrik sebagai Alat Analisis

Dalam analisis data penelitian-penelitian sosial saat ini sering digunakan Statistika Nonparametrik. Statistika ini termasuk dalam ketegori Statistika Inferensial, yang dipakai untuk menafsirkan **parameter** (populasi) berdasarkan **statistik** (sampel) melalui pengujian statistik atau yang lebih dikenal dengan Uji Signifikansi.



GAMBAR 3.2 *Pengolahan Data Statistik Parametrik dan NonParametrik*

Beberapa hal yang perlu diperhatikan sebelum menggunakan Statistika Nonparametrik antara lain:

1. Penggunaan Statistika Nonparametrik *hanyalah untuk data penelitian yang berasal dari sampel*, sebab jika data penelitian berasal dari populasi (sensus) hasil pengukurannya berupa parameter, dengan demikian tidak perlu ditafsirkan lagi tetapi bisa langsung diinterpretasikan.
2. Statistika Nonparametrik mensyaratkan pengambilan data dengan cara random, karena di dalamnya mengandung kaidah-kaidah probabilitas.
3. Perhatikan hipotesis penelitian, karena hipotesis tersebut mengindikasikan apakah pengujian (uji signifikansi) harus dilakukan satu sisi atau dua sisi.
4. Perhatikan dengan cermat, apakah penelitian kita terdiri atas kasus satu sampel, dua sampel, atau lebih dari dua sampel.
5. Jika penelitian merupakan kasus dua sampel atau lebih, perhatikan dengan lebih teliti, apakah merupakan sampel yang berpasangan atau tidak berpasangan.

Beberapa pengujian nonparametrik berikut akan dikelompokkan berdasarkan sampel penelitian, dan tersedia dalam paket *software* SPSS (*Statistical Package for Social Sciences*) yang banyak digunakan dalam penelitian sosial dengan cara operasi yang relatif mudah.

1. ***Kasus Satu Sampel:*** Misalnya kita ingin melakukan penelitian untuk meneliti apakah betul sekolah-sekolah favorit telah secara adil memberi kesempatan kepada pria dan wanita, atau kepada semua masyarakat dari berbagai tingkat ekonomi. Uji signifikansi yang bisa digunakan antara lain:
    a. *Uji Binomial:* Digunakan untuk menguji perbedaan proporsi sebuah populasi, jika data berskala nominal dan hanya memiliki dua kategori.
    b. *Uji Chi-Kuadrat Sampel Tunggal:* Digunakan untuk menguji perbedaan proporsi sebuah populasi, jika data berskala nominal dan memiliki lebih dari dua kategori.
    c. *Uji Kolmogorov-Smirnov Sampel Tunggal:* Digunakan untuk menguji perbedaan proporsi sebuah populasi, jika data berskala ordinal.
2. ***Kasus Dua Sampel Berpasangan:*** Misalnya kita ingin melakukan penelitian prestasi atau perilaku siswa sebelum dan setelah dilakukan perubahan kurikulum. Jadi sampel yang sama diukur dua kali, pertama dilakukan pengukuran terhadap prestasi atau perilaku sebelum perubahan kurikulum, dan kedua pengukuran prestasi atau perilaku siswa dilakukan setelah perubahan kurikulum. Uji signifikansi yang bisa digunakan antara lain:
    a. *Uji Mc-Nemar:* Digunakan untuk menguji perbedaan proporsi dua populasi yang berpasangan, jika data berskala nominal dan hanya memiliki dua kategori.
    b. *Uji Tanda:* Digunakan untuk menguji perbedaan nilai tengah ranking dua populasi yang berpasangan, jika data berskala ordinal.
    c. *Uji Tanda Wilcoxon:* Digunakan untuk menguji perbedaan nilai tengah ranking dua populasi yang berpasangan *dengan lebih halus*, jika data berskala ordinal.
3. ***Kasus Dua Sampel Tidak Berpasangan:*** Misalnya kita ingin melakukan penelitian prestasi atau perilaku siswa antara dua sekolah yang berbeda atau antara dua kota yang berbeda atau antara sekolah di pedesaan dan perkotaan. Dengan demikian untuk masing-masing sampel hanya diukur

satu kali, tetapi dengan model pengukuran yang sama. Uji signifikansi yang bisa digunakan antara lain:

a. *Uji Chi-Kuadrat Dua Sampel Berpasangan:* Digunakan untuk menguji perbedaan proporsi dua populasi yang tidak berpasangan, jika data berskala nominal dengan dua atau lebih dari dua ketegori.
b. *Uji U Mann-Whitney:* Digunakan untuk menguji perbedaan nilai tengah ranking dua populasi yang tidak berpasangan, jika data berskala ordinal.
c. *Uji Kolmogorov-Smirnov Dua Sampel:* Digunakan untuk menguji "sembarang" perbedaan (median, dispersi, dan skewness) dua populasi yang tidak berpasangan, jika data berskala ordinal.

4. ***Pengukuran Korelasi dan Uji Signifikansinya:*** Dalam sebuah penelitian kadang kala kita ingin mengetahui apakah ada hubungan antara peubah satu dengan yang lainnya, untuk keperluan tersebut sering digunakan pengukuran korelasi. Besarnya koefisien korelasi (r), serta arah dari koefisien (negatif atau positif) dapat dipakai sebagai indikasi kuat tidaknya hubungan antara dua buah peubah serta bagaimana arah hubungannnya.

Hal yang perlu dipahami dalam penggunaan ukuran korelasi adalah, bahwa koefisien korelasi yang dihasilkan tidak otomatis menunjukkan bahwa peubah yang satu berpengaruh terhadap peubah lain, tetapi hanya menunjukkan tingkat asosiasi kuat lemahnya hubungan, sementara penentuan peubah indpenden dan dipenden ditentukan berdasarkan teori.

Jika pengukuran korelasi didasarkan pada sampel, koefisien korelasi adalah *statistik*, untuk menjawab apakah angka korelasi tersebut berlaku juga dalam populasinya sebagai *parameter*, perlu dilakukan pengujian signifikansi. Kalau berdasarkan hasil pengujian angkanya signifikan, maka koefisien korelasi sebagai *statistik* bisa diebut sama dengan *parameter*-nya.

Pengukuran korelasi yang biasa digunakan dalam penelitian sosial aantara lain:

a. *Koefisien Kontingensi (C):* Digunakan untuk mengukur keeratan hubungan antara dua peubah yang berskala nominal. Misalnya apakah ada hubungan antara proporsi jenis kelamin murid SMA dengan proporsi keinginan mereka untuk melanjutkan pendidikan ke perguruan tinggi pada jurusan eksata dan non eksata.

b. *Koefisien Korelasi Rank Kendall (τ):* Digunakan untuk mengukur keeratan hubungan antara dua peubah yang berskala ordinal. Misalnya apakah ada hubungan antara ranking test masuk SMP dengan dengan ranking di semester pertama kelas 1 SMP.
c. *Koefisien Korelasi Rank Sperman (r_s):* Digunakan untuk mengukur keeratan hubungan antara dua peubah yang berskala ordinal. Pengukuran korelasi ini lebih banyak digunakan karena metodenya yang lebih sederhana.

Bab 4
Penyajian Data

4.1 PENGANTAR

Dalam bab ini akan diuraikan tentang penyajian data dalam statistika. Data yang dikumpulkan dari lokasi penelitian, pada umumnya belum teratur, dan masih merupakan bahan keterangan yang sifatnya kasar dan data mentah. Salah satu tugas statistik adalah menyusun data mentah dan menyajikannya dengan cara yang teratur, ringkas dan mudah dimengerti, sehingga dengan jelas dapat memberikan gambaran yang tepat mengenai ciri atau makna yang terkandung dalam data tersebut. Untuk itu statistik mempersiapkan dua bentuk penyajian data yaitu berbentuk Tabel dan Grafik atau Diagram. Bentuk penyajian ini bergantung kepada jenis data dan keperluannya. Secara ringkasnya, **Tabel** merupakan kumpulan angka-angka yang disusun sedemikian rupa menurut kategori-kategori tertentu, sehingga kumpulan angka-angka tersebut (data) mudah untuk dianalisis, sedangkan **Grafik (diagram)** merupakan gambar yang menunjukkan secara visual data berupa angka, yang biasanya juga berasal dari tabel-tabel yang telah dibuat.

4.2 TABEL

Penyajian tabel digunakan sebagai pilihan yang sering dipakai oleh peneliti atau penyaji informasi. Pengolahan data untuk keperluan analisis awal atau analisis lanjutan akan lebih baik apabila di sajikan terlebih dahulu dalam tabel yang baik. Tabel adalah alat penyajian data angka dalam bentuk baris-baris dan kolom-kolom. Data angka yang dikumpulkan disusun dan didistribusikan ke dalam baris-baris dan kolom-kolom menurut klasifikasi datanya. Misalnya, jumlah pegawai diklasifikasikan menurut jenis kelamin, umur dan pendidikan. Jumlah penduduk diklasifikasikan menurut suku dan agama.

Jenis tabel berdasarkan penyajian terdiri dari dua jenis, yaitu tabel referensi dan tabel ikhtiar. **Tabel referensi** memiliki fungsi sebagai "gudang keterangan" karena tabel tersebut menyajikan keterangan yang rinci dan disusun secara khusus untuk kepentingan referensi. **Misal**, tabel-tabel yang terdapat dalam laporan sensus umumnya merupakan tabel yang memberikan keterangan secara umum bagi kepentingan referensi. Seringkali tabel semacam ini disebut tabel umum (*general table*),

Tabel ikhtisar atau juga dinamakan tabel naskah (*text table*), umumnya berbentuk singkat, sederhana dan mudah dimengerti. Fungsi tabel ikhtisar adalah memberi gambaran yang sistematis tentang peristiwa-peristiwa yang merupakan hasil penelitian atau observasi. Tabel ikhtisar dapat berasal dari tabel referensi atau dari beberapa table ikhtisar yang lainnya. **Tabel ikhtisar banyak digunakan** dalam penulisan laporan perusahaan maupun tulisan ilmiah. Salah satu jenis tabel ikhtisar adalah tabel yang isinya menggambarkan perbandingan. Angka yang perbandingkan tentu saja diletakkan dalam kolom yang berdampingan. Jika angka-angka absolut yang diperbandingkan terlalu besar, maka dapat disajikan dalam bentuk rasio atau persentase untuk lebih memudahkan. Stressing atau penekanan hal-hal yang dianggap penting dapat dilakukan dengan cara meletakkan angka-angka dalam kolom yang berada di sisi kiri dan yang tidak diberi penekanan diletakkan dalam kolom yang berada di sisi kanannya.

4.2.1 Penyusunan Data dalam Bentuk Tabel

Data yang diperoleh secara acak harus diolah terlebih dahulu untuk memudahkan proses analisis. Adapun langkah-langkah dalam penyusunan tabel supaya mudah dibaca adalah sebagai berikut:

a. Penyusunan secara alfabetis
   Tabel ini menyajikan data berdasarkan kolom nama kompartemen yang disusun secara alfabetis dimulai dari alfabet paling awal yang ada dalam kolom tersebut (ascending).
b. Penyusunan secara geografis
   Tabel ini menyajikan data berdasarkan kolom nama kompartemen yang disusun secara geografis dimulai dari lokasi paling barat, misalnya di Indonesia adalah provinsi Banda Aceh.
c. Penyusunan menurut besaran angka-angka
   Tabel ini menyajikan data berdasarkan kolom yang diberikan penekanan dan disusun menurut besarnya angka-angka, baik dari kecil ke besar (ascending) maupun dari besar ke kecil (descending).
d. Penyusunan secara historis
   Data yang disajikan dalam tabel diklasifikasikan secara kronologis atau historis, biasanya dimulai dari waktu yang paling dahulu atau paling lama.
e. Penyusunan atas dasar kelas-kelas yang lazim
   Penyajian data dalam tabel dimana penyusunan pos-pos keterangan dalam kompartemen tabel dilakukan berdasarkan kelas-kelas yang umum digunakan dalam dunia statistik. Misalnya Impor, seringkali dibagi ke dalam tiga kategori ekonomi, yaitu: a. barang konsumsi, b. bahan mentah serta bahan pelengkap, dan c. barang modal.
f. Penyusunan secara progresif
   Pada tabel ini, penyusunan pos-pos keterangan dalam kompartemen tabel harus dilakukan sedemikian rupa agar angka akhir dari tiap pos harus merupakan hasil perkembangan angka-angka yang telah ada sebelumnya. Cara penyusunan yang digunakan dalam menyusun pos-pos keterangan dalam kompartemen tabel harus diusahakan agar tabel referensi disusun untuk tujuan referensi, sedangkan tabel ikhtisar disusun untuk tujuan perbandingan serta penekanan pada pospos yang dianggap penting oleh penyusun.

Setelah proses penyusunan data ke dalam tabel selesai, Tabel yang telah disajikan dapat diklasifikasikan ke dalam 4 bentuk tabel, yaitu:

1. Tabel Klasifikasi Acak/Tunggal
2. Tabel Klasifikasi Berkelompok

3. Tabel Kontingensi
4. Tabel Frekuensi/Distribusi Frekuensi

Untuk lebih memudahkan dan memperjelas struktur masing-masing ke empat bentuk tabel di atas, maka akan disajikan dalam tabel berikut ini:

a) Tabel Klasifikasi Acak

TABEL 4.1 ***Jumlah Wisata Mancanegara yang berkunjung ke Bali Mengikut Kawasan (2020).***

Kawasan	Banyaknya Wisman
Afrika	10.758
Amerika	79.010
Asean	100.967
Oseania	244.227
Asia Non Asean	334.247
Eropa	300.264
Jumlah	1 069.473

Sumber: BPS-Provinsi Bali, 2020

b) Tabel Klasifikasi Berkelompok

TABEL 4.2 ***Jumlah Buruh Pabrik "ABC" berdasarkan Jenis Kelamin dan Pendidikan Tahun 2019***

Jenis Kelamin	Pendidikan			Jumlah
	SMP	SMA	PT	
Laki-laki	35	72	15	122
Perempuan	42	21	10	73
Jumlah	77	93	25	1196

Sumber: Data Karangan

TABEL 4.3 ***Jumlah Aparatur Sipil Negara Menurut Provinsi dan Jenis Kelamin Pada Tahun 2016***

Provinsi	Jenis Kelamin		Jumlah
	Laki-laki	Perempuan	
Pusat	547.625	370.819	918.444
Provinsi	166.213	135.568	301.781
Kabupaten/Kota	1.503.655	1.650.469	3.154.124
Jumlah	2.217.493	2.156.856	4.374.349

Sumber: BPS- Badan Kepegawaian Negara (BKN), 2017

c) Tabel Kontingensi

TABEL 4.4 ***Hubungan pekerjaan dan jenis olahraga yang diminati***

Pekerjaan	Kebiasaan Merokok			Jumlah
	Sepak Bola	Basket	Renang	
Dokter	7	4	6	8
Guru	5	1	7	10
Pegawai Swasta	1	6	5	15
Pegawai Negeri	3	5	3	9
Jumlah	15	13	14	42

Sumber: Data Karangan

d) Tabel Distribusi Frekuensi

TABEL 4.5 ***Penghasilan Nasabah Mitra Multi Finance Tahun 2018***

Gaji Bulanan (Juta Rupiah)	Frekuensi (f)
0,5 - 1,0	10
1,1 - 1,5	13
1,6 - 2,0	8
2,1 - 2,5	11
2,6 - 3,0	9
3,1 - 3,5	7
3,6 - 4,0	3
Jumlah	61

Sumber: Data Karangan

TABEL 4.6 ***Nilai Ujian Akhir Sekolah di SMA IT-NR semester Genap tahun ajaran 2018/2019***

Nilai Statistik Ekonomi	Banyak Mahasiswa (f)
11 – 20	2
21 – 30	4
31 – 40	7
41 – 50	12
51 – 60	30
61 – 70	10
71 – 80	5
81 – 90	5
Jumlah	70

Sumber: Data Hipotetis

4.3 GRAFIK

4.3.1 Fungsi Grafik dalam Statistik

Grafik tidak lain adalah alat penyajian data yang tertuang dalam bentuk lukisan, baik lukisan garis, gambar maupun lambang. Jadi, dalam penyajian data angka melalui grafik, angka itu disajikan dalam bentuk lukisan garis, gambar, atau lambang tertentu. Data statistik dapat disajikan dalam bentuk tabel atau grafik. Penyajian data dalam bentuk grafik umumnya lebih menarik perhatian dan mengesankan. Penyajian data statistik secara grafis mempunyai berbagai fungsi. Grafik atau diagram seringkali digunakan dalam iklan dengan maksud agar konsumen memperoleh kesan yang mendalam terhadap ciri-ciri produk yang diiklankan. Kegiatan produksi lebih mudah dilihat dan dipelajari secara visual bila dinyatakan dalam angka-angka dan digambarkan secara grafis. Peta pengawasan kualitas merupakan alat yang penting dalam melakukan pengawasan produk maupun pengawasan proses produksi.

Grafik penjualan suatu perusahaan memberi gambaran yang sederhana dan menarik mengenai perkembangan hasil penjualan yang telah dicapai oleh perusahaan yang bersangkutan. Pada hakekatnya grafik dan tabel seyogyanya digunakan secara bersama-sama. Grafik statistik lebih mudah dan menarik dibanding tabel statistik. Selain itu, grafik dapat melukiskan suatu peristiwa secara lebih mengesankan dan tidak membosankan. Namun demikian, penyajian secara grafis hanyalah bersifat aprosimatif. Angka-angka yang pasti dan rinci tentang suatu peristiwa dimuat dalam tabel. Oleh karena itu, analisis dan interpretasi data umumnya dilakukan terhadap data yang terdapat dalam tabel statistik. Dengan kata lain, data angka divisualisasikan. Dibandingkan dengan tabel, grafik memiliki keunggulan sebagai berikut:

a. Penyajian data melalui grafik tampak lebih menarik
b. Grafik dapat dengan cara lebih cepat memperlihatkan gambaran umum menyeluruh tentang sesuatu perkembangan perubahan maupun perbandingan.
c. Grafik yang dibuat menurut aturan yang tepat dan benar, akan terasa lebih jelas dan lebih dimengerti pembaca.

Sebaliknya, grafik memiliki kelemahan dibandingkan dengan tabel, antara lain:

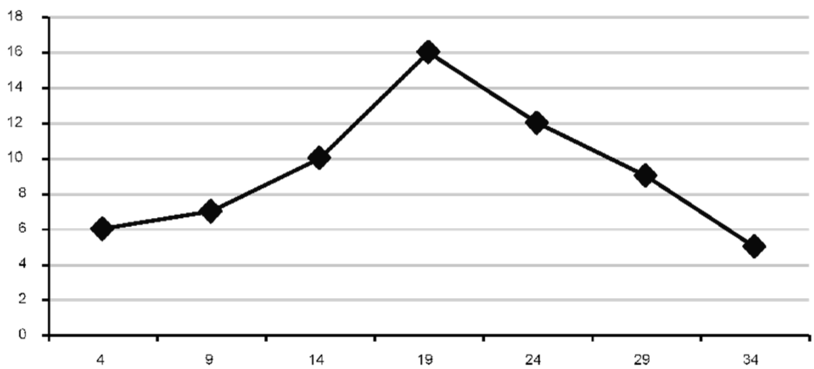
a. Membuat grafik jauh lebih sukar dan memakan waktu, biaya atau alat yang lebih banyak.
b. Data yang dapat disajikan dalam bentuk grafik, sangat terbatas. Jika data yang akan disajikan banyak macamnya, maka lukisan grafiknya menjadi ruwet dan memusingkan.
c. Umumnya grafik bersifat kurang teliti. Dalam tabel, dapat dimuat angka sampai tingkat ketelitian yang setinggi-tingginya. Misalnya, angka 6.35, 7.25 dapat dimuat dalam tabel, namun tidak mungkin dilakukan pada grafik.

Dengan mengetahui kelemahan dan keunggulan tabel dan grafik sebagai alat penyajian data, maka hal itu dapat dijadikan pedoman untuk menetapkan apakah data yang sedang dihadapi akan disajikan melalui tabel atau grafik.

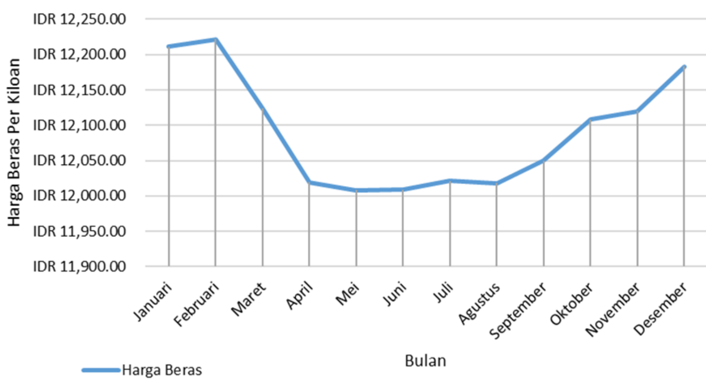
4.3.2 Jenis Diagram dalam Statistik

a. Diagram Garis (Polygon)

Diagram garis sering disebut juga peta garis (line chart) atau kurva (curve), merupakan bentuk penyajian yang paling banyak dipakai dalam berbagai laporan perusahaan maupun penelitian ilmiah. Data statistik dapat diklasifikasikan atas ciri-ciri kronologis, geografis, kuantitatif maupun kualitatif. Salah satu bentuk data yang dapat diklasifikasi secara kronologis adalah data deret berkala (time series). Sebagian besar distribusi data dapat diklasifikasi secara kuantitatif dalam bentuk distribusi frekuensi. Hasil kedua cara klasifikasi tersebut dapat digambarkan secara visual dalam bentuk kurva. Sedangkan data yang diklasifikasikann berdasarkan geografis maupun kualitatif, jarang digambarkan dalam bentuk kurva. Data demikian dapat digambarkan dengan peta balok (bar chart) atau bentuk peta lainnya.



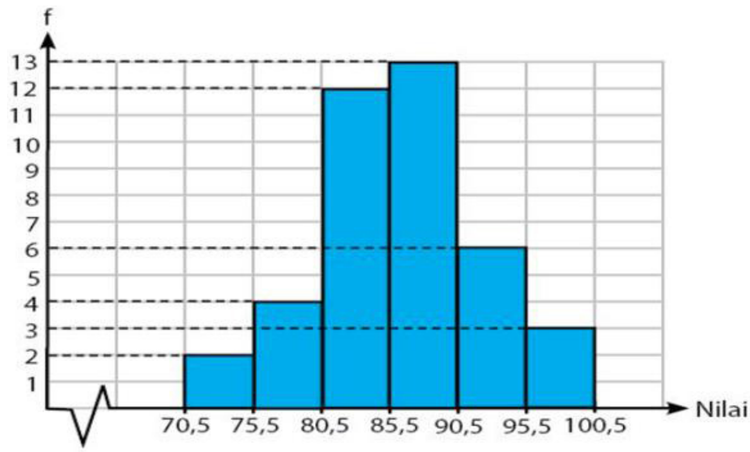
GAMBAR 4.1 *Poligon Data Acak*



GAMBAR 4.2 *Rata-rata Harga Beras Per Bulan di Tingkat Perdagangan Besar/ Grosir Indonesia (Rupiah/Kg) Tahun 2019*

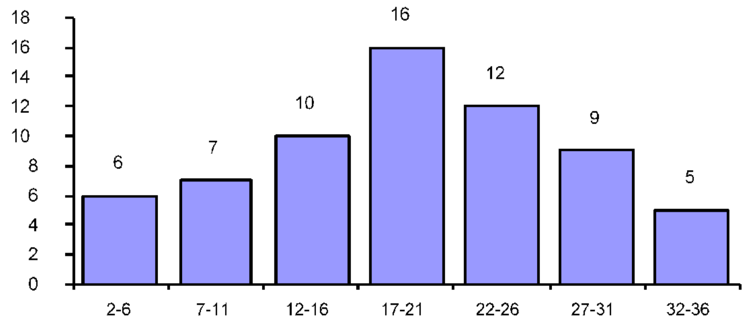
b. Diagram Batang (Histogram)

Grafik histogram adalah grafik yang tersusun dari segi empatsegi empat yang didirikan pada absis, membentang selebar-lebarnya kelas. Tinggi dari segi empat itu sebanding dengan frekuensi masing-masing kelas yang diwakili. Seperti halnya dengan poligon, ordinatnya juga menyatakan frekuensi dan absisnya menyatakan tingkatan-tingkatan gejala. Perbedaannya, absis pada poligon, dinyatakan dengan kelas, sedangkan absis pada histogram dinyatakan dengan batas nyatanya. Histogram dapat dibuat dari distribusi frekuensi data tunggal maupun dari distribusi frekuensi data kelompok.



GAMBAR 4.3 ***Data nilai Matematika 40 siswa***

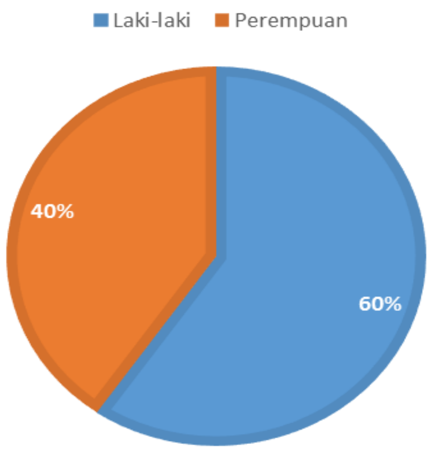
Diagram Batang ini digunakan untuk memahami persoalan secara visual. Dalam diagram batang, lebar kelas diambil dari selang kelas distribusi frekuensi, sedangkan frekuensi masing-masing kelas ditunjukkan oleh tinggi batang. Diagram histogram berbeda dengan diagram batang dalam hal lebar, yaitu batang digunakan batas kelas dan bukan limit kelasnya. Ini untuk menghilangkan jeda antar batang sehingga antar batang memberikan kesan ”padat”.



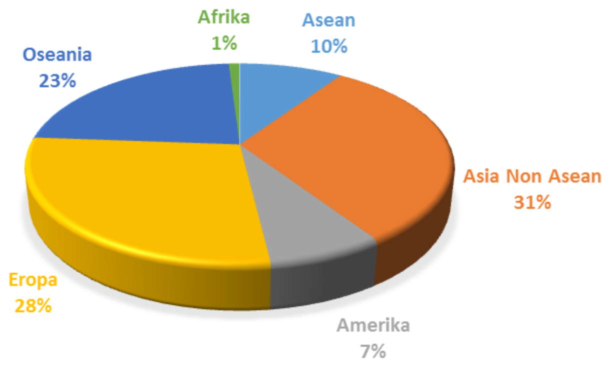
GAMBAR 4.4 *Data lamanya pegawai bekerja di PT. XYZ*

c. Diagram Lingkaran (*Pie Diagram*)

Cara lain untuk menyajikan data hasil penelitian adalah dengan diagram lingkaran (pie chart). Diagram lingkaran digunakan untuk membandingkan data dari berbagai kelompok. Diagram lingkaran biasanya digunakan untuk menyatakan perbandingan jika data terdiri atas beberapa kelompok atau kategori.



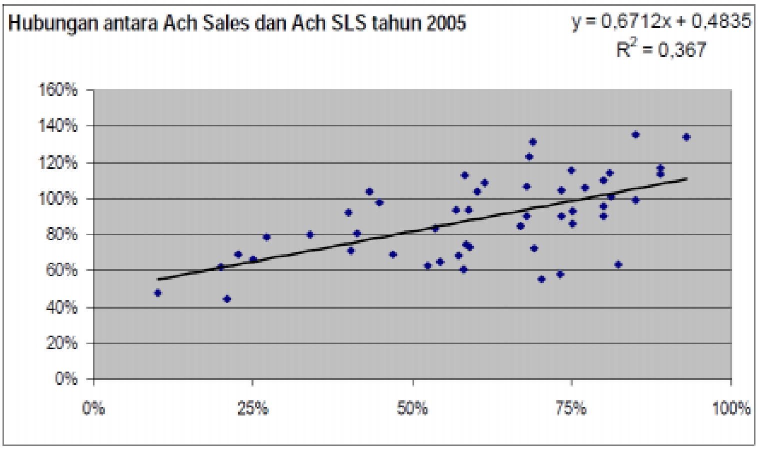
GAMBAR 4.5 *Persentase Aparatur Sipil Negara di Pusat Pemerintahan Tahun 2016*



GAMBAR 4.6 *Persentase Jumlah Wisata Mancanegara yang Datang Langsung ke Bali Menurut Kawasan Tahun 2020*

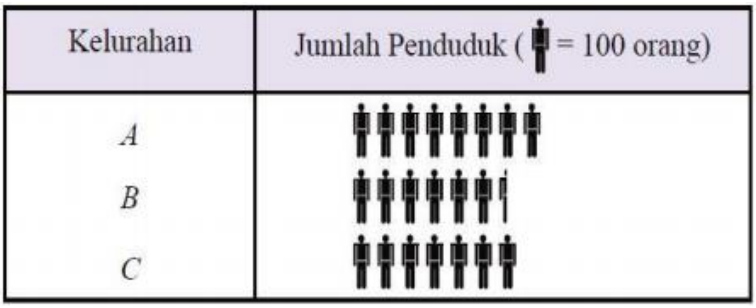
d. Diagram Pencar

Diagram pencar atau disebut juga dengan diagram titik (diagram sebaran) ialah diagram yang menunjukkan gugusan titiktitik setelah garis koordinat sebagai penghubung diputus. Untuk kumpulan data yang terdiri dari dua peubah dengan nilai kuantitatif maka diagramnya dapat dibuat dalam sistem sumbu koordinat dan gambarnya akan merupakan kumpulan titik-titik yang terpencar.



e. Diagram gambar/lambang

Diagram gambar atau diagram lambang sering dipakai untuk mendapatkan gambaran kasar sesuatu hal dan sebagai alat visual. Diagram gambar sangat menarik untuk dilihat, lebih-lebih jika gambar atau lambang yang digunakan cukup baik dan menarik.Setiap satuan jumlah tertentu dibuat sebuah gambar atau lambang sesuai dengan macam datanya, misalnya untuk dapat mengenai jiwa, penduduk dan pegawai dibuat gambar orang, satu gambar mewakili data sebanyak 1000 orang. Untuk data bangunan, gedung sekolah maka dibuat gambar gedung yang satu gambarnya mewakili 10 gedung. Produksi mobil pertahun, maka dibuat gambar mobil di mana satu gambar mobil mewakili 1000 mobil yang diproduksi dalam satu tahun.



f. Diagram peta (kartogram)

Diagram peta dalam pembuatannya digunakan peta geografis tempat data terjadi. Dengan demikian diagram ini melukiskan keadaan dihubungkan dengan tempat kejadiannya. Salah satu contoh yang sudah terkenal ialah jika membuka buku peta bumi. Di dalamnya terdapat peta daerah atau pulau dengan mencantumkan gambar gunung, gambar padi atau palawija yang mengilustrasikan daerah produksinya dan sebagainya.

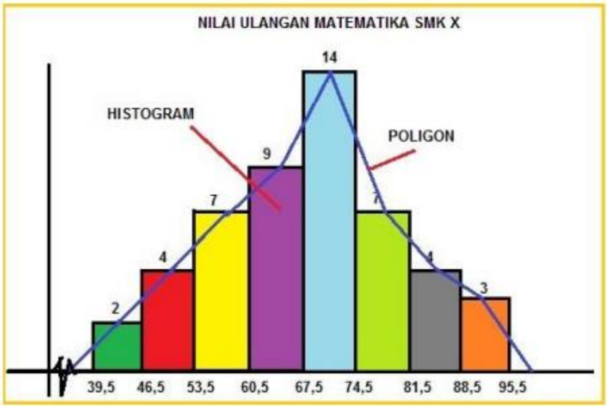
Di samping bentuk penyajian data di atas itu, masih ada bentuk pelaporan lainnya dari hasil kerja statistik, yaitu bentuk perumusan, atau dalam statistik lebih sering kita kenal dengan nama bentuk “tekstular”. Bentuk ini secara teratur selalu mengikuti semua penyajian data statistik yang sudah dianalisa dan disimpulkan. Kesimpulan-kesimpulan statistik biasanya dirumuskan dalam katakata atau kalimat-kalimat. Kerap kali kerja statistik hanya menghasilkan konklusi-konklusi yang dirumuskan dalam kata-kata tanpa disertai penyajian dalam bentuk lainnya. Juga tidak jarang penyajian hasil kerja statistik diberikan dalam bentuk yang bermacam-macam sekaligus.

g. Poligon frekuensi

Poligon frekuensi adalah grafik garis dari data dalam tabel distribusi frekuensi yang menghubungkan frekuensi setiap nilai tengah interval kelas yang dimulai dari interval kelas ke nol (sebelum interval kelas pertama) sampai dengan interval kelas ke n + 1 (sesudah interval kelas ke n). Frekuensi interval kelas ke nol sama dengan nol, dan demikian pula halnya frekuensi interval kelas ke n + 1 pun sama dengan nol.

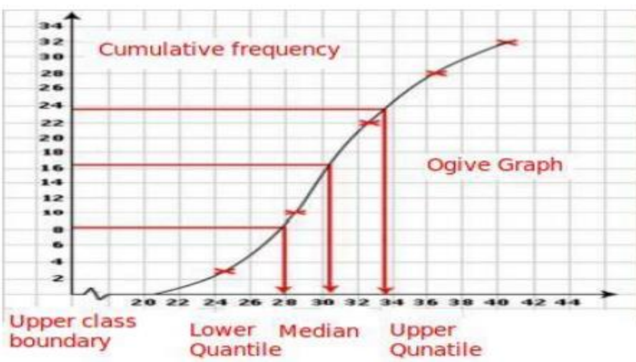
Pada dasarnya pembuatan grafik poligon sama dengan histogram, hanya cara membuat batas-batas pada sumbu mendatar(horizontal) yang berbeda. Perbedaan antara poligon dan histogram yaitu:

a. Histogram menggunakan nilai tepi kelas dalam menentukan absis (batas-batas sumbu horizontalnya) sedangkan poligon menggunakan nilai titik tengah sebagai absis (batas-batas sumbu horizontalnya).
b. Histogram berwujud berupa segi empat sedangkan grafik poligon berwujud garis-gris atau kurva yang saling berhubungan satu dengan lainnya yang ujung awal dan akhirnya menutup pada sumbu horizontal.



h. Ogive

Ogive adalah grafik garis dari suatu data dalam distribusi frekuensi kumulatif dengan nilai-nilai skala horizontalnya berupa nilai tepi kelas (batas kelas) setiap interval kelas dan nilai skala vertikalnya berupa frekuensi kumulatif. Jadi dalam nilai-nilai skala pada absis (horizontal) grafik ogive sama dengan pada absis histogram, sementara bentuk grafiknya berupa grafik garis seperti pada poligon. Ada dua macam ogive yaitu ogive dengan frekuensi kumulatif kurang dari atau sama dengan dan ogive dengan frekuensi kumulatif lebih dari atau sama dengan.



4.3.3 Beberapa Peraturan Umum Tentang Menggambar Grafik

a) Pemilihan jenis grafik
Jenis grafik statistik yang akan disajikan oleh pembuat laporan harus dipilih agar dapat menyajikan gambaran mengenai suatu data secara efektif bagi pembaca. Jika dilihat dari fungsinya, setiap jenis grafik statistik memiliki kelebihan-kelebihan khusus. Namun demikian, grafik yang baik harus bersifat sederhana dan jelas. Grafik yang rumit biasanya disajikan untuk orang yang sangat mengerti permasalahan atau yang sangat mahir dalam ilmu statistik. Pemilihan jenis grafik yang akan disajikan oleh pembuat laporan tidak dapat semata-mata diserahkan pada kebijakan penggambar grafik, kecuali bila pembuat laporan yakin bahwa penggambar memiliki pengetahuan yang baik tentang data yang disajikan, tujuan penyajian, dan kemampuan pembaca dalam menarik kesimpulan dari grafik.

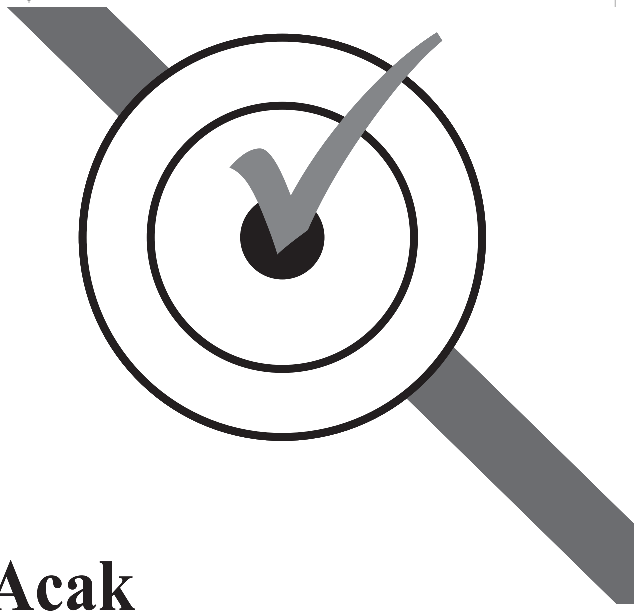
b) Nama (titel), skala sumbu, sumber dan catatan
Kegunaan serta pengaturan nama, sumber dan catatan dalam sebuah tabel berlaku juga untuk grafik statistik. Nama grafik dapat diletakkan di atas atau di bawah gambar grafik. Meski demikian banyak statistisi berpendapat bahwa peletakan nama di atas grafik akan lebih efektif jika dibandingkan dengan di bawah grafik. Skala *horizontal* dan *vertical* dalam peta garis, diagram kolom, dan peta balok sebenarnya memiliki kesamaan dalam arti dengan nama kolom dan kompartemen dalam tabel statistik.

c) Skala dan garis kisi-kisi
Jarak yang sama pada skala grafik sebenarnya menyatakan jarak nilai yang sama pula. Nilai skala bertujuan memberi gambaran yang aproksimatif tentang jumlah kuantitatif, sedangkan jumlah yang eksak dan rinci secara seksama harus dibaca dari tabel statistiknya.

Garis kisi-kisi harus digambarkan secara lebih tipis dari pada garis skalanya. Peta garis umumnya memiliki garis kisi-kisi baik yang bersifat mendatar maupun vertikal. Peta kolom hanya membutuhkan garis kisi-kisi yang mendatar. Peta balok mendatar membutuhkan garis kisi-kisi vertikal. Pada beberapa penyajian grafik, garis kisi-kisi demikian dapat juga tidak digambarkan sama sekali atau hanya digambarkan secara sebagian saja.

d) Pemberian tekanan pada penggambaran grafik

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 terputusputus, garis padat (solid line) atau garis tebal. Garis padat lebih memberi tekanan dari pada garis terputus-putus, sedangkan garis tebal lebih menarik perhatian dari pada garis yang tipis.



Bab 5

Analisis Data Acak

Pada Bab sebelumnya kita sudah membahas materi tentang pengumpulan, pengolahan dan penyajian data dalam analisis statistik. Sehingga, kita lebih mudah mengenal apa itu sampel, populasi dan bagaimana menyajikan data-data tersebut sesuai dengan kaedahnya. Dalam bab ini akan lebih banyak fokus pembahasan kepada analisis data acak atau data tunggal. Data acak dapat dianalisis dengan menggunakan tiga ukuran yaitu ukuran pemusatan, ukuran letak dan ukuran penyebaran. Ukuran pemusatan meliputi ukuran untuk nilai rata-rata (Mean), nilai tengah (Median), dan nilai yang sering muncul (Modus). Selanjutnya, ukuran letak adalah ukuran yang dapat mengetahui letak dari pada pembagian data sesuai kelompok masing-masing fraktil. Fraktil terdiri dari kuartil, desil dan persentil. Ukuran terakhir yaitu ukuran penyebaran merupakan ukuran yang dapat membandingkan sifat yang dimiliki oleh kelompok data. Perbandingan ini dapat dilihat melalui nilai simpangan baku yang diperoleh. Apabila nilai hitung rata-rata sebuah data adalah sama besar maka kelompok data yang mempunyai nilai simpangan baku yang lebih kecil lebih mendekat ke pusat data berbanding dengan kelompok data dengna nilai simpangan baku yang lebih besar.

5.1 UKURAN PEMUSATAN

Secara definisi ukuran pemusatan adalah ukuran yang menunjukkan pusat dari sekumpulan data. Pusat dari pada data tersebut akan menunjukkan nilai rata-rata dari seluruh rangkaian data yang tersaji. Nilai rata-rata ini dikenal juga dengan sebutan ukuran tendensi pusat (measure of central tendency). Dalam statistik, ukuran pemusatan dapat dihitung melalui lima perhitungan antara lain (1) rata-rata hitung (mean), (2) modus, (3) median, (4) rata-rata ukur, dan (5) rata-rata harmonis.

5.1.1 Nilai Rata-Rata *Mean*

Rata-rata hitung dari data tunggal dapat diperoleh dengan cara menjumlahkan seluruh nilai dan membaginya dengan banyaknya data. Jika x_1, x_2, ..., x_n merupakan nilai-nilai data dengan jumlah data sebanyak n, rata-ratanya adalah:

$$\boxed{\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}} \quad \text{atau} \quad \boxed{\bar{x} = \frac{\sum x_i}{n}}$$

Keterangan:

$\bar{x}$ = Rata-rata (baca: " *x* bar");

n = Banyaknya data;

$\sum x_i$ =Jumlah seluruh data.

Contoh:

Hitunglah rataan dari data: 7, 5, 9, 4, 8, 6, 10, 7!

Jawab:

$$\bar{x} = \frac{7 + 5 + 9 + 4 + 8 + 6 + 10 + 7}{8} = \frac{56}{8} = 7$$

Jadi, rataan dari data 7, 5, 9, 4, 8, 6, 10, 7 adalah 7.

Contoh:
Hitunglah rataan dari data: 8, 10, 9, 7, 5, 4, 9, 5, 5, 4, 3, 6, 7!

Jawab:

$$\bar{x} = \frac{8 + 10 + 9 + 7 + 5 + 4 + 9 + 5 + 5 + 4 + 3 + 6 + 7}{13} = \frac{82}{13} = 6.31$$

Jadi, rataan dari data 8, 10, 9, 7, 5, 4, 9, 5, 5, 4, 3, 6, 7 adalah 6.31
Jadi, rata-rata nilai pada deretan angka di atas adalah 6.31.

5.1.2 Nilai Tengah *Median*

Median merupakan nilai tengah dari sekelompok data yang nilai tiap observasi telah disusun dari yang terkecil ke terbesar. Tidak sensitif terhadap nilai ekstrim. Median digunakan untuk mengukur pemusatan kalau distribusi mencong (*skewed*) secara jelas. Dapat dihitung pada distribusi yang tidak komplit sekalipun, misalnya distribusi yang berakhir terbuka (contoh 150-169; 170-189; 190-209; 210+).

1) Bila jumlah observasi (=n) ganjil, maka median adalah nilai observasi ke: $\frac{n+1}{2}$ dari urutan nilai observasi kecil ke besar.

   Contoh: 5, 4, 5, 6, 7, 1, 5, 3, 4, 6, 9. Tentukan median
   Urutkan data: 1, 3, 4, 4, 5, 5, 5, 6, 6, 7, 9
   Median(M_e) = $\frac{n+1}{2} = \frac{11+1}{2} = 6$

2) Bila banyaknya observasi (=n) genap, maka median adalah nilai di antara observasi ke: $\frac{n}{2}$ dan $\frac{n}{2} + 1$, diambil rata-rata.

   Contoh: 1, 3, 4, 4, 5, 5, 5, 6, 6, 7

$$n = 10 \longrightarrow \frac{n}{2} = 5 \; dan \; 5 + 1 = 6$$

$$\downarrow$$

$$M_e = \frac{5 + 6}{2} = 5,5$$

Contoh:
Peneliti bank ingin melihat jumlah median para sepuluh nasabah bank umum. Diambil data secara acak dan diperoleh angka seperti berikut:

7	3	2	9	11	6	25	8	10	12

Angka di atas dalam juta rupiah. Tentukan nilai median nasabah tersebut.

Jawaban:

(1) Mengurutkan nila angka dari terkecil ke terbesar

2	3	6	7	8	9	10	11	12	25

(2) Letak Median, untuk n = 10 (genap)
L_{Md} antara data yang ke $\frac{10}{2}$ dan $\frac{10+2}{2}$ yaitu antara data urutan yang ke-5 dan ke-6

(3) Nilai Median (Md) = nilai data urutan ke-5 ditambah ke-6 dibagi 2

$$nilai\ \ data\ = \frac{x_5 + x_6}{2} = \frac{8+9}{2} = 8.5$$

Jadi, Median nilai tabungan sepuluh nasabah tersebut adalah Rp 8,5 juta.

5.1.3 Modus Nilai Sering Muncul *Mode*

Modus merupakan nilai yang paling sering muncul (frekuensi terbesar) dari seperangkat data atau observasi. Modus dari serangkaian data adalah nilai (atau sifat) yang paling banyak terjadi, atau sifat/keadaan yang frekuensinya terbesar. Untuk data kuantitatif modus menunjukkan nilai yang paling banyak muncul dan untuk data kualitatif modus menunjukkan sifat atau keadaan yang paling banyak terjadi. Dengan demikian serangkaian data, mungkin tidak mempunyai modus, satu modus, dua modus atau lebih. Modus juga mencerminkan yang paling tipikal atau kasus yang paling umum. Kalau kita ingin segera mengetahui nilai pemusatan, maka kita menghitung modus. Seperangkat data dapat saja tidak memiliki modus, tetapi sebaliknya dapat pula memiliki beberapa modus. Kalau satu modus saja disebut unimodal, dua modus disebut bimodal dan kalau tanpa modus disebut nonmodal.
Modus (*crude mode*) = nilai yang paling sering muncul

Contoh: 1, 3, 4, 4, 5, 5, 5, 6, 6, 7

$M_0 = 5$

Bila sekumpulan data belum disusun dalam distribusi frekuensi, maka cara menghitung/menentukan modusnya adalah sebagai berikut:

(1) Hitung frekuensi masing-masing data atau sifat atau keadaan
(2) Menentukan modusnya. Data yang frekuensinya terbesar (untuk data kuantitatif) atau sifat/keadaan yang paling sering terjadi (untuk data kualitatif) merupakan **modusnya**

Contoh:
Nilai simpanan tabungan sebuah sampel acak yang berjumlah 15, berasal dari nasabah bank disajikan sebagai berikut:

100	50	100	60	300	250	75	80
50	100	100	75	250	100	300	

Hitunglah modusnya dan berikan interpretasi

Jawaban

TABEL 5.1 ***Cara Menghitung Modus Nilai Simpanan Tabungan 15 Nasabah***

Nilai Tabungan (x_i)	Banyak/Frekuesi (f)
50	2
60	1
75	2
80	1
100	5
250	2
300	2
Jumlah	15

Berdasakrn Tabel 5.1 di atas nilai frekuensi terbesar = 5.

Data (x_i) dengan urut 5 (x_5= 100) memiliki frekuensi tebesar yaitu 5 (f = 5). Jadi, modusnya adalah tabungan dengan nilai 100 juta rupiah. Modus = 100 juta rupiah, memiliki arti bahwa nilai tabungan deposan yang paling banyak ada (untuk sampel terpilih) yaitu tabungan yang bernilai 100 juta rupiah.

Contoh:

1. Untuk data yang tidak dikelompokkan

   Di bawah ini adalah data sepuluh pasien yang menjalani perawatan di rumah sakit:

Pasien ke	**Lama perawatan (hari)**	**Pasien ke**	**Lama perawatan (hari)**
1	28	6	11
2	14	7	29
3	14	8	22
4	24	9	18
5	14	10	14

Hitung: rerata, median, modul lama perawatan dari pasien-pasien ini!

a. Rata-rata

$$\bar{x} = \frac{\sum x}{n} = \frac{11 + 14 + 14 \ldots + 24 + 28 + 29}{10}$$

$$\bar{x} = \frac{188}{10} = 18{,}8 \text{ hari}$$

b. Median

Urutan nilai observasi adalah sebagai berikut:

11; 14; 14; 14; 14; 18; 22; 24; 28; 29

Karena banyaknya observasi genap, maka median merupakan rata-rata nilai dari observasi ke $\frac{n}{2} = \frac{10}{2} = 5$ dan $\frac{n}{2} + 1 = 6$

Jadi: Median = $\frac{14+18}{2}$ = 16 hari

c. Modus

Oleh karena 14 hari adalah nilai yang paling sering muncul, maka modus adalah 14 hari.

5.1.4 Rata-Rata Ukur

Rata-rata ukur (**GM**) serangkaian n data adalah akar pangkat n dari hasil kali nilai - nilai seluruh data tersebut. Dengan catatan salah satu nilai - nilai data tesebut tidak negatif atau nol. Rata-rata ukur (*Geometric Mean*), terutama digunakan untuk: (1) menghitung rata-rata data rasio, seperti rata- rata persen, rata- rata nilai indeks dan rata-rata nilai relatif, dan (2) untuk menghitung rata-rata laju perubahan. Seperti rata-rata pertumbuhan penduduk, rata-rata pertumbuhan ekonomi, rata- rata perubahan indeks ekonomi, rata- rata kenaikan penjualan, rata-rata kenaikan produksi dari satu periode ke periode waktu lain. Jadi, rata-rata ukur (GM) dapat dirumuskan sebagai berikut:

$$GM = \sqrt[n]{(X_1)(X_2)(X_3)\ \dots\dots.(X_n)}$$

GM = rata-rata ukur
x_1 = nilai data yang pertama,
x_n = nilai data yang ke-n
n = banyaknya data/banyak pengamatan

Contoh:
Data menunjukkan Indeks harga penjualan minyak makan selama 4 tahun seperti berikut:

200, 125, 100, dan 150

Hitung nilai rata-rata indeks harga penjualan minyak makan tersebut.

Jawaban

$x_1 = 100$, $x_2 = 125$, $x_3 = 150$, dan $x_4 = 200$

$$GM = \sqrt[4]{(X_1)(X_2)(X_3)(X_4)}$$

$$GM = \sqrt[5]{(100)(125)(150)(200)}$$

$$GM = 139.15$$

Jadi, rata-rata indeks harga penjualan beras tersebut 139,15.

5.1.5 Rata-rata Harmonis

Rata-rata harmonis (*Harmonic Mean*) adalah banyaknya data atau pengamatan (n) dibagi dengan jumlah kebalikan nilai datanya. Dalam prakteknya rata-rata harmonis banyak digunakan untuk mencari rata-rata nilai data yang berbeda untuk sejumlah pengamatan yang sama. Misalnya menghitung kecepatan rata-rata, untuk jarak yang sama dengan kecepatan yang berbeda, menghitung harga rata-rata per unit dalam beberapa kali pembelian dengan sejumlah uang yang sama (tetap) yang harga per unitnya berbeda dalam setiap pembelian.

Rata-rata Harmonis Data yang tidak berkelompok Rata-rata harmonisnya dapat dihitung dengan rumus:

$$H = \frac{\text{n}}{\sum \frac{1}{X_1}}$$

H = rata-rata harmonis,

n = banyaknya data/pengamatan

x_i = nilai data yang ke-i

Contoh:

Seorang pedagang menyediakan anggaran sebesar Rp 600.000,00 tiap bulannya dalam jangka waktu 6 bulan untuk pembelian daging sapi. Jika harga per kg daging sapi mulai bulan pertama sampai dengan bulan yang ke enam sebagai berikut: Rp 20.000,00; Rp 25.000,00; Rp 30.000,00; Rp 40.000,00; Rp 50.000,00; Rp 60.000,00. Tentukanlah harga rata-rata tiap kg daging sapi tersebut

Jawaban

Disini banyak pengamatan (frekuensi fembelian) adalah 6 kali, jadi n = 6. Nilai pengamatan yaitu: x1 = 20.000, x2 = 25.000, x3 = 30.000, x4 = 40.000, x5= 50.000, dan x6= 60.000.

Dengan menggunakan rumus di atas, diperoleh hasil sebagai berikut:

$$H = \frac{\mathrm{n}}{\sum \frac{1}{X_1}}$$

$$H = \frac{6}{\frac{1}{20.000} + \frac{1}{25.000} + \frac{1}{30.000} + \frac{1}{40.000} + \frac{1}{50.000} + \frac{1}{60.000}}$$

$$H = \frac{6}{\frac{30}{60.000} + \frac{24}{60.000} + \frac{20}{60.000} + \frac{15}{60.000} + \frac{12}{60.000} + \frac{10}{60.000}}$$

$$H = \frac{6}{\frac{111}{60.000}}$$

$$H = 32.432{,}40$$

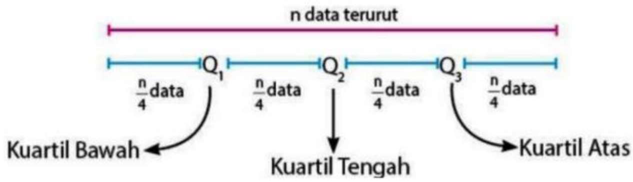
Jadi, harga rata-rata (harmonis) per kg daging sapi tersebut adalah Rp 32.432, 40

5.2 UKURAN LETAK

Kuartil adalah nilai-nilai yang membagi data atas empat bagian yang sama banyaknya setelah data tersebut diurutkan. Ada tiga buah kuartil, yaitu kuartil pertama atau kuartil bawah (Q 1), kuartil kedua/tengah (Q2/Me) dan kuartil ketiga atau kuartil atas (Q3). Agar kita dapat mengetahui lebih jauh mengenai karakteristik data observasi dengan beberapa ukuran sentral, kita sebaiknya mengetahui beberapa ukuran lain, yaitu ukuran letak. Ada empat macam ukuran letak yang akan di bahas pada bagian ini, yaitu Median, Kuartil, Desil, dan Persentil. Median dalam ukuran letak adalah sebuah nilai yang membagi serangkaian data atau suatu distribusi menjadi dua bagian yang sama, yaitu lima puluh persen (50%) dari keseluruhan data nilainya terletak di bawah (nilai) median dan lima puluh persen (50%) lagi nilainya terletak di atas (nilai) median.

5.2.1 Kuartil

Kuartil adalah ukuran letak yang membagi data observasi menjadi empat bagian yang sama banyak. Oleh karena itu masing-masing bagian mengandung 25% data observasi. Pada satu set data observasi mempunyai tiga buah kuartil, yaitu Q_1, Q_2, Q_3 seperti yang ditunjukkan dalam gambar berikut:



GAMBAR 5.1 *Letak Kuartil berdasarkan nilai Kuartil Bawah, Tengah dan Atas*

Untuk menentukan nilai kuartil data observasi yang tidak berkelompok (*ungrouped data*) melalui langkah-langkah sebagai berikut ini:

1) Urutkan data observasi dari kecil ke besar
2) Tentukan letak kuartilnya
   Menentukan letak Q_1, Q_2, Q_3 dapat digunakan formulasi sebagai berikut:

   Letak $Q_1 = \frac{n+1}{4}$

   Letak $Q_2 = \frac{2(n+1)}{4}$

   Letak $Q_3 = \frac{3(n+1)}{4}$
3) Tentukan nilai kuartilnya.
   Nilai Q_1, Q_2, Q_3 adalah data observasi yang terletak pada letak Q_1, Q_2, Q_3.

Contoh:

Berikut ini adalah data mengenai nilai 7 orang peserta ujian Statistik di UMB Yogyakarta:

78 56 66 48 80 70 76

Tentukan Q_1, Q_2, Q_3 Jawab:

Untuk menentukan Q_1, Q_2, Q_3 maka langkah-langkah yang digunakan adalah sebagai berikut:

Contoh:

Urutkan nilai tersebut dari kecil ke besar

48 56 66 70 76 78 80

Tentukan letak Q_1, Q_2, Q_3 dengan formula

Letak $Q_1 = \frac{7+1}{4} = 2$

Letak $Q_2 = \frac{2(7+1)}{4} = 4$

Letak $Q_3 = \frac{3(7+1)}{4} = 6$

Jadi letak Q_1 pada urutan data ke 2, letak Q_2 pada urutan data ke 4, dan letak Q_3 pada urutan data ke 6.

Contoh:

Tentukan nilai Q_1, Q_2, Q_3

No urut	1	2	3	4	5	6	7
Nilai	48	56	66	70	76	78	80

Q_1 Q_2 Q_3

Nilai Q_2 adalah juga merupakan median dari nilai peserta ujian tersebut. Apabila banyaknya data observasi menunjukkan bilangan genap, maka median terletak diantara dua nomor urut.

Contoh:

Tentukan nilai kuartil bawah (Q_1, kuartil kedua (Q_2/Me), dan kuartil atas (Q_3) pada data berikut:

a. 5, 4, 2, 10, 14, 12, 11;
b. 5, 9, 7, 4, 13, 10;

Jawaban

a 5,4,2,10,14,12,11

Data diurutkan menjadi:

2, 4, 5, 10, 11, 12, 14,

$Q_1 = 4$ Q_2/Me $Q_3 = 12$

Langkah pertama mencari Q_2/Me terlebih dahulu, kemudian Q_1, yaitu nilai-nilai yang ≤Q2, selanjutnya Q_3 yang nilai-nilainya $\geq Q_2$. Jadi, nilai $Q_1 = 4$; Q_2/Me = 10; dan $Q_3 = 12$.

b. 5,9,7,4,13,10

Data diurutkan menjadi:

4, 5, 7, 9, 10, 13,

Q_1 Q_2 Q_3

Contoh:

Nilai deposito (juta rupiah) dari sampel acak sebelas deposan telah disusun sebagai berikut: 18 19 20 23 24 25 27 30 32 35 dan 36.

Tentukanlah

(a) Q_1 dan berikanlah interpretasi.

(b) Q_2 dan berikanlah interpretasi.

Jawaban

Oleh karena datanya telah diurut dari nilai yang terkecil sampai dengan nilai yang terbesar, maka langsung dapat ditentukan letak K_1 dan K_2 sebagai berikut:

(a) Q_1 = ,,,?

Letak Q_1

n = 11, x = 1

$$Q_1 = \frac{n+1}{4} = \frac{11+1}{4}$$

$$Q_1 = 3$$

Letak kuartil pertama (Q_1), yaitu pada data dengan urutan ketiga. Nilai Q_1 sama dengan nilai data urutan ketiga, yaitu Rp 20 juta.

Interpretasi nilai Q_1. Nilai Q_1 = Rp 20 juta, memiliki arti kurang-lebih 25% dari seluruh (sampel) deposan tersebut memiliki deposito yang nilainya lebih kecil dari Rp 20 juta, dan sisanya lagi yaitu 75% memiliki deposito dengan nilai lebih besar dari Rp 20 juta.

(b) Q_2 = ...?

Letak Q_2

n = 11, dan x = 2

$$Q_2 = \frac{2(n+1)}{4} = \frac{2(11+1)}{4}$$

Letak kuartil kedua (Q_2), yaitu pada data dengan urutan ke-6 Nilai Q_2 sama dengan nilai data urutan ke-6 yaitu Rp 25 juta.

Interpretasi nilai Q_2. Nilai Q_2 = Rp 25 juta, artinya kurang-lebih 50% dari seluruh (sampel) deposan tersebut memiliki deposito yang nilainya kurang dari atau paling tinggi Rp 25 juta, dan 50% nya lagi memiliki deposito dengan nilai lebih besar dari atau paling rendah Rp 25 juta.

5.2.2 Desil

Desil adalah ukuran letak yang membagi data observasi menjadi sepuluh bagian yang sama banyak. Oleh karena itu masing-masing bagian mengandung 10% data observasi. Pada satu set data observasi mempunyai sembilan buah desil, yaitu $D_1, D_2, \ldots, D_9$. Untuk data dengan jumlah genap maka hasil hitung pada nilai desil yang diperoleh akan berbentuk angka pecahan sehingga letak desil pada kumpulan data berjumlah genap diantara kedua data bilangan bulat.

Untuk menghitung nilai desil pada data acak atau tunggal maka rumus untuk memperoleh nilai desil jika banyak data n dan Di adalah desil ke-i adalah sebagai berikut:

$$Di = \frac{i(n+1)}{10}$$

Dengan:

D_i = letak desil ke-x

i = 1, 2, 3, 4, …, 9

n = Banyaknya data

Contoh:

Tentukan D_3, dan D_5 dari; 6, 4, 6, 4, 7, 5, 6, 5, 8, 7, 7, 7, 8, 6!

Jawaban

Data diurutkan menjadi; 4, 4, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8

Data	4	4	5	5	6	6	6	6	7	7	7	7	8	8
Data ke-	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Letak D_i = data ke $\frac{i(n+1)}{10}$

Letak D_3 = data ke- $\frac{3(14+1)}{10}$

= data ke- 4,5 ($X_{4,5}$)

Dengan interpolasi diperoleh:

$$D_3 = X_4 + 0{,}5\ (X_5 - X_4)$$
$$= 5 + 0{,}5(6 - 5)$$
$$= 5{,}5$$

Letak D_5 = data ke- $\frac{5(14+1)}{10}$ = data ke- 7,5 ($X_{7,5}$)

Dengan interpolasi diperoleh:

$$D_5 = X_7 + 0{,}5\,(X_8 - X_7)$$

$$D_5 = 6 + 0{,}5(6 - 6)$$

$$D_5 = 6$$

Contoh:

Harga per lembar saham (ribu rupiah) dari 29 perusahaan yang diambil sebagai sampel acak, disajikan sebagai berikut:

15	15	16	20	21	22	23	24	27	28	30	30	31	31	
32	33	34	35	36	36	37	37	38	39	39	40	41	43	43

Hitunglah:

(a) D_1 dan berikanlah interpretasi

(b) D_7 dan berikanlah interpretasi

Jawaban

(a) Menghitung dan memberikan interpretasi terhadap D_1

Data tersebut telah disusun dari nilai terkecil sampai nilai terbesar.

Letak D_1: x = 1, dan n = 29

$$Di = \frac{1(29 + 1)}{10}$$

$$Di = 3$$

Letak desil pertama (D_1), yaitu pada data dengan urutan ke-3. Nilai D_1 sama dengan nilai data dengan urutan ke-3 yaitu Rp 16 ribu.

Interpretasi. Nilai D_1 = Rp 16 ribu artinya kurang-lebih 10% dari keseluruhan sampel yang berupa saham tersebut harga per lembarnya kurang dari Rp 16 ribu dan 90% nya lagi harga per lembarnya lebih dari Rp16 ribu

(b) Menghitung dan memberikan interpretasi terhadap D_7

Letak D_7: x = 7, dan n = 29

$$Di = \frac{7(29+1)}{10} = \frac{210}{10} = 21$$

Nilai D_7 sama dengan nilai data dengan urutan ke -21 yaitu Rp 37 ribu. Interpretasi. Nilai D_7= Rp 37 ribu artinya kurang-lebih 70% dari keseluruhan sampel harga per lembar nya kurang dari Rp 37 ribu dan 30% nya harga per lembarnya lebih dari Rp 37 ribu.

5.2.3 Persentil

Persentil adalah ukuran letak yang membagi data observasi menjadi seratus bagian yang sama besar. Oleh karena itu masing-masing bagian mengandung 1 % data observasi. Pada satu set data observasi mempunyai 99 persentil, yaitu: P_1, P_2,…,P_{99}.

Untuk menghitung nilai persentil pada data acak atau tunggal maka tahapan perhitungannya adalah sebagai berikut:

(1) Menyusun data dari nilai terkecil sampai yang terbesar atau sebaliknya.

(2) Menentukan letak persentil. Letak persentil ke-x ditentukan dengan rumus:

$$P_i = \frac{i(n+1)}{100}$$

Dengan,

P_i = data ke- $\left(\frac{i(n+1)}{100}\right)$,

i = 1,2,3,……99,

n = banyak data/pengamatan

(3) Menghitung nilai Persentil ke-x (P_x). Bila $L_P x$ merupakan bilangan utuh (bulat positif) Maka, nilai persentil ke-x (P_x) = nilai data dengan urutan ke $\frac{x(n+1)}{100}$.

Contoh:

Tentukan P_{30}, dan P_{75} dari; 6, 4, 6, 4, 7, 5, 6, 5, 9, 7, 10, 7, 10, 6!

Jawaban

Data diurutkan menjadi; 4, 4, 5, 5, 6, 6, 6, 6, 7, 7, 7, 9, 10, 10

Data	4	4	5	5	6	6	6	6	7	7	7	9	10	10
Data ke-	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Letak P_i = data ke $\frac{i(n+1)}{100}$

Letak P_{30} = data ke- $\frac{30(14+1)}{100}$

= data ke- 4,5 ($X_{4,5}$)

Dengan interpolasi diperoleh:

$$P_{30} = X_4 + 0{,}5(X_5 - X_4)$$
$$= 5 + 0{,}25(6 - 5)$$
$$= 5{,}25$$

Letak P_{75} = data ke- $\frac{75(14+1)}{100}$

= data ke- 11,25 ($X_{11,25}$)

Dengan interpolasi diperoleh:

$$P_{75} = X_{11} + 0{,}5(X_{12} - X_{11})$$
$$= 7 + 0{,}25(9 - 7)$$
$$= 7{,}5$$

Contoh:

Tingkat hunian (dalam %) sembilan hotel bintang 5 yang disurvei baru-baru ini, hasilnya sebagai berikut:

85 70 65 82 65 60 90 80 75

Hitunghlah P_{30} dan berikan interpretasi.

Jawaban

Susun terlebih dahulu data tersebut dari nilai terkecil sampai nilai terbesar, sebagai berikut:

60 65 65 70 75 80 82 85 90

i = 30, dan n = 9

$$P_i = \frac{30\,(9+1)}{100} = \frac{30\,(9+1)}{100} = 3$$

Maka, Nilai P_{30} sama dengan nilai data urutan ketiga yaitu 65%. Nilai P_{30} = 65%, artinya kurang-lebih 30% dari seluruh data (sembilan tingkat hunian tersebut) nilainya kurang dari 65% dan 70% nya lagi nilainya lebih dari 65%.

5.3 UKURAN PENYEBARAN

5.3.1 Pengertian Tentang Ukuran Penyebaran

Digunakan untuk menunjukkan keadaan berikut:

a. Gambaran variabilitas data

   Yang dimaksud dengan variabilitas data adalah suatu ukuran yang menunjukkan besar kecilnya perbedaan data dari rata-ratanya. Ukuran ini dapat juga disebutkan sebagai ukuran yang menunjukkan perbedaan antara data satu dengan yang lainnya. Ukuran pemusatan (Mean, Median, dan Modus) ini dapat kita gunakan untuk menggambarkan keadaan sekumpulan data, tetapi gambaran itu masih kurang lengkap apabila tidak disertai dengan ukuran-ukuran penyebaran. Hal ini disebabkan karena dengan ukuran gejala pusat saja mungkin beberapa kumpulan data sebenarnya berbeda dapat disimpulkan sama.

b. Perbedaan nilai satu observasi terhadap nilai observasi lainnya

   Rata-rata dari serangkaian nilai-nilai observasi tidak dapat diinterpretasikan secara terpisah dengan dispersi (sebaran) nilai-nilai tersebut terhadap rata-ratanya. Jika terdapat keseragaman/kesamaan nilainilai observasi, X_i, maka dispersi nilai-nilai tersebut akan sama dengan nol, dan rata-ratanya akan sama dengan nilai X_i. Semakin besar variasi nilai-nilai X_i, maka rata-rata distribusi semakin kurang representatif.

Contoh:

Tabel 5.1 Rata-rata hitung hasil test mata kuliah statistik deskriptif kelompok A dan B.

Kelompok	Hasil test					
A	60	65	50	60	65	60
B	65	90	50	70	60	60

Mahasiswa A: $\overline{X}$ = 360/6 = 60
Mahasiswa B: $\overline{X}$ = 360/6 = 60

Rata-rata hasil test kedua mahasiswa tersebut tidak berbeda, namun dispersi hasil test mahasiswa B (30 sampai dengan 90) jauh lebih besar dari pada varisasi hasil test mahasiswa A (50 sampai dengan 65). Hal ini berarti hasil test mahasiswa A jauh lebih konsisten (stabil) disbanding mahasiswa B. Tingkat dispersi berhubungan erat dengan sifat kesamaan/kesejenisan data. Misalnya data tentang besarnya modal pedagang kaki lima khusus makanan, akan kecil variasinya jika dibandingkan dengan data seluruh pedagang kaki lima tanpa melihat jenis dagangannya. Secara umum, suatu rata-rata akan cukup representatif bagi serangkaian nilai-nilai observasi bila nilai-nilai tersebut diperoleh dari data yang bersifat sejenis bagi tujuan pengamatan tertentu.

5.3.2 Pengukuran Jarak (Range)

Pengukuran jarak sebuah distribusi merupakan pengukuran disperse yang paling sederhana. Jarak sebuah distribusi frekuensi dirumuskan sebagai “selisih atau beda antara pengukuran nilai terbesar dan nilai terkecil yang terdapat dalam sebuah distribusi frekuensi”. Atau secara matematis dapat ditulis sebagai berikut:

$$R = X_u - X_i$$

Keterangan:
R = Range data observasi
X_u = Nilai tertinggi
X_i = Nilai terendah

Beberapa Catatan Tentang Pengukuran dan Penggunaan Jarak

1) Hasil pengukuran jarak (range) sebenarnya sudah dapat menggambarkan disperse (variasi) nilai-nilai observasi dengan cara yang paling sederhana.

Jika kita ingin memperoleh hasil pengukuran dispersi secara kasar dan cepat, maka ukuran range dapat digunakan.

2) Range bukan merupakan pengukuran dispersi distribusi yang memuaskan karena hasil pengukurannya jelas tergantung pada kedua nilai ekstrim tanpa mengikutsertakan pola dispersi nilai-nilai observasi secara keseluruhan.

Contoh:
Berikut ini adalah nilai ulangan harian 10 siswa mata pelajaran statistika di SMA Mercu Buana Yogyakarta:

56 66 78 94 48 82 50 76 80 70

Range nilai 10 siswa yang ikut ulangan harian statistika tersebut dapat ditentukan dengan menggunakan formula:

$R = X_u - X_i$

R = 94 – 48 =46

Contoh:
Aset delapan (miliar rupiah) perusahaan yang dipetik dari laporan keuangan nya yang telah dipublikasikan melalui media cetak (surat kabar) nasional adalah sebagai berikut:

252 250 265 270 275 280 301 352

Hitunglah rangenya.

Jawaban

$R = X_u - X_i$

R = 352 - 252

R = 100

Jadi, range-nya Rp 100 miliar

5.3.3 Pengukuran Deviasi Kuartil

Nilai-nilai yang ordinatnya membagi seluruh distribusi dalam 4 (empat) bagian yang sama dinamakan nilai-nilai kuartil. Q1 merupakan kuartil pertama, Q2 merupakan kuartil kedua dan sama dengan median (Q2 = md), sedangkan Q3 dinamakan kuartil ketiga. Dalam distribusi kuartil, 50% dari semua nilai-nilai observasi seharusnya terletak antara Q1 dan Q3. Jarak antara Q1 dan Q3 dinamakan jarak inter-kuartil (inter-quartilrange). Makin kecil jarak tersebut, maka makin tinggi tingkat konsentrasi distribusi tengah seluas 50% dari seluruh distribusi.

Secara teoritis, pengukuran deviasi kuartil sebuah sampel data baik data acak maupun data berkelompok dapat rumuskan sebagai:

$$dQ = \frac{Q_3 - Q_1}{2}$$

Selanjutnya dapat dikatakan bahwa deviasi kuartil adalah sebesar +dQ atau –dQ dari mediannya.

Pada dasarnya, pengukuran deviasi kuartil sama seperti pengukuran jarak (*range*). Pengukurannya didasarkan pada jarak antara Q1 dan Q3. Pengukuran tersebut tidak dipengaruhi oleh dispersi dari seluruh nilai-nilai observasi, deviasi kuartil hanya mengikutsertakan dispersi nilai-nilsi observasi X_i yang didistribusikan di tengah-tengah seluruh distribusi seluas 50% saja.

5.3.4 Pengukuran Deviasi Rata-rata (*Mean Deviation*)

Deviasi rata-rata untuk data acak dapat menggunakan rumus di bawah ini. Dispersi serangkaian nilai-nilai observasi akan kecil bila nilai-nilai tersebut berkonsentrasi sekitar rata-ratanya. Sebaliknya, dispersinya akan besar bila nilai-nilai observasi tersebar jauh dari rata-ratanya. Deviasi rata-rata dari seluruh nilai-nilai observasi dapat dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum(x_i - \bar{x})}{n}$$

Sedangkan pengukuran deviasi atas dasar nilai-nilai absolut dapat dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum|x_i - \bar{x}|}{n}$$

Contoh:
Carilah deviasi rata-rata data berikut ini:

40	50	70	55
55	72	66	60
60	54	85	65
45	67	80	75
70	80	55	80

Jawaban
Dimana i=1,2,3,4,…..,20

$$d_{\bar{x}} = \frac{\sum|x_i - \bar{x}|}{n} = \frac{706}{20} = 35{,}31$$

Contoh:
Pengeluaran per minggu (ratus ribu rupiah) 5 orang ibu rumah tangga (A, B, C, D dan E) untuk keperluan biaya hidup (dalam ratus ribu rupiah) pada tahun 2012, adalah sebagai berikut:

12 16 18 20 24

Tentukanlah deviasi rata-ratanya

Jawaban

TABEL 5.2 ***Deviasi Rata-rata Pengeluaran 5 Rumah Tangga (Ratus Ribu Rupiah)***

Ibu Rumah Tangga	Pengeluaran Bulanan (x_i)	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$
A	12	-6	6
B	16	-2	2
C	18	0	0
D	20	2	2
E	24	6	6
Jumlah	90	0	16

Dari Tabel 5.2 di atas, diketahui n = 5, x_i = 90 dan $|x_i - \bar{x}|$ = 16

$$d_{\bar{x}} = \frac{\sum(x_i - \bar{x})}{n}$$

$$d_{\bar{x}} = \frac{90}{5}$$

$$d_{\bar{x}} = 18$$

Selanjutnya rumus $d_{\bar{x}}$ dihitung dan didapat seperti berikut:

$$d_{\bar{x}} = \frac{\sum|x_i - \bar{x}|}{n}$$

$$d_{\bar{x}} = \frac{16}{5}$$

$$d_{\bar{x}} = 3{,}2$$

Jadi, deviasi rata-rata biaya hidup per minggu lima ibu rumah tangga adalah Rp 320 ribu. Nilai AD = Rp 320 ribu memiliki arti bawa secara rata-rata kelima pengeluaran biaya hidup tersebut menyimpang sebesar Rp 320 ribu dari nilai rata-ratanya.

Contoh:
Suku bunga deposito berjangka 3 bulan (% per tahun) untuk enam valuta asing yang ditawarkan oleh sebuah bank, dicatat sebagai berikut:

Valuta Asing	**Tingkat suku bunga (% per tahun)**
AUS $	6,50
Pound	6,50
Yen	3,00
Sin $	3,50
DM	5,50
HK $	4,50

Dengan menganggap data tersebut sampel acak,
(a) Hitunglah variansinya
(b) Hitunglah deviasi standar keenam suku bunga valuta asing tersebut. Berikan interpretasi terhadap nilainya.

Jawaban
(a) Menghitung variansi keenam suku bunga tersebut
Tabel Cara Menghitung Variansi dan Simpangan Baku Suku Bunga Keenam Valuta Asing

Valuta Asing	**Suku bunga (x_i)**	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert^2$
AUS$	6,50	1,58	2,50
Pound	6,50	1,58	2,50
Yen	3,00	- 1,92	3,69
Sin $	3,50	- 1,42	2,02
DM	5,50	0,58	0,34
HK $	4,50	- 0,42	0,18
Jumlah	29,50		11,23

Dari tabel 6.3 dapat diketahui bahwa

N= 6,30, $\sum x_i = 29{,}5 \; dan \; d_{\bar{x}} = \sum |\mathrm{xi} - \bar{x}|^2 = 11{,}23$

$\bar{x}$ dihitung dan didapat sebagai berikut:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{29{,}5}{6} = 4{,}92$$

Selanjutnya per rumus s^2 didapat:

$$d_{\bar{x}} = \frac{\sum |\text{xi} - \bar{x}|^2}{n-1} = \frac{11{,}23}{5} = 2{,}25$$

Jadi, variansinya adalah 2,25 %

(b) Menghitung Deviasi Standar

Per rumus (6.9) s didapat,

$$s = \frac{\sqrt{\sum |\text{xi} - \bar{x}|^2}}{n-1} = \sqrt{s^2}$$

$$s = \sqrt{2.25}$$

$$s = 1.50$$

Jadi, simpangan baku keenam suku bunga tersebut adalah 1,50% per tahun.

Interpretasi nilai s. Nilai s = 1,50% artinya bahwa rata-rata penyimpangan kelima suku bunga tersebut dari rata-ratanya sebesar 1,50%

5.3.5 Pengukuran Varians dan Deviasi Standar

Varians digunakan untuk melihat kehomogenan data secara kasar, dimana nilai hasil perhitungan varians sebagai titik pusat dari penyebaran data.

Contoh:

Seorang guru matematika melakukan tes prestasi dengan membagi siswa dalam 3 kelompok, yaitu A,B, dan C. Dalam satu kelompok terdapat 5 siswa. Walaupun dibentuk kelompok namun untuk tes dikerjakan secara individu. Didapat hasil sebagai berikut:

KELOMPOK	NILAI					X
A	50	50	50	50	50	50
B	60	40	50	55	45	50
C	30	70	90	10	50	50

a. Varians dan deviasi standar dari data yang belum dikelompokkan

Karl Pearson merumuskan pengukuran **varians** sebagai:

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

Standarisasi unit-unit pengukuran di atas dilakukan melalui proses pengakaran, dan dinamakan **deviasi standar**, sebagai berikut:

$$s = \sqrt{s^2}$$

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

b. Varians dan deviasi standar dari data acak

- Rumus Fisher dan Wilks

  Varians dari Fisher dan Wilks:

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

- Deviasi standar dari Fisher dan Wilks:

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

- Varians dan deviasi standar **populasi**

  Varians polupasi:

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

- Deviasi standar populasi:

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

Bab 6
Analisis Data Berkelompok

6.1 PENGANTAR

Jika pada bab sebelumnya fokus utama materi adalah menjelaskan dan menjabarkan tentang analisis pada data acak maka pada bab ini akan menitikberatkan pada analisis data berkelompok. Perbedaan antara kedua jenis data tersebut ada pada kelas intervalnya. Data acak merupakan sekumpulan data sederhana yang belum disusun mengikut kelas intervalnya. Sehingga data ini lebih dikenal dengan sebutan data acak atau data tunggal. Berbeda dengan data berkelompok yang di mana data sudah dikelompokkan berdasarkan kelas interval data dan menghasilkan kelas frekuensi data. Oleh karena itu, ciri pembeda dari kedua data adalah kelas interval data tersebut.

6.2 UKURAN PEMUSATAN

6.2.1 Nilai Rata-Rata *Mean*

Data berkelompok adalah data yang telah mempunyai distribusi frekuensi. Dalam ukuran pemusatan pada nilai rata-rata data berkelompok pada dasarnya tidak mempunyai banyak perbedaan berbanding data acak. Jika pada data acak nilai *X* merupakan rata-rata untuk masing-masing individu yang diwakili, maka pada data berkelompok nilai *X* adalah titik tengah dari kelas interval data. Untuk memperoleh nilai rata-rata pada data berkelompok maka dapat menggunakan rumus sebagai berikut:

$$\bar{x} = \frac{\sum f.xi}{\sum f}$$

Dimana:

$\bar{x}$ = Titik tengah (mid point) kelas interval ke I

xi = Titik tengah interval kelas

f = Frekwensi observasi pada kelas interval ke i

f_x = Jumlahkan frekwensi tiap kelas interval

Contoh:

Dalam suatu kelas yang mengikuti ulangan Babasa Indonesia, diperoleh data: siswa yang mendapat nilai 5 ada 6 orang, nilai 6 ada 12 orang, nilai 7 ada 15 orang, nilai 8 ada 7 orang. Tentukan rataannya!

Jawaban

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{6(5) + (12)6 + 15(7) + 7(8)}{6 + 12 + 15 + 7} = \frac{263}{40} = 6{,}575$$

Jadi, rataanya adalah 6,575

Contoh:

Tabel data tinggi badan mahasiswa FKIP UMB- Yogyakarta diambil 50 mahasiswa secara random. Hasil Pengukuran tinggi badan adalah sebagai berikut

Interval Kelas	f_i
164,5 – 167,5	6
167,5 – 170,5	7
170,5 – 173,5	8

173,5 – 176,5	11
176,5 – 179,5	7
179,5 – 182,5	6
182,5 – 185,5	5
Jumlah	50

Jawaban

Interval Kelas	F	Xi	f*xi
164,5 – 167,5	6	166	996
167,5 – 170,5	7	169	1183
170,5 – 173,5	8	172	1376
173,5 – 176,5	11	175	1925
176,5 – 179,5	7	178	1246
179,5 – 182,5	6	181	1086
182,5 – 185,5	5	184	920
Jumlah	50		8732

Maka,

$$\overline{x} = \frac{\sum f.xi}{\sum f} = \frac{8732}{50} = 174{,}64$$

6.2.2. Nilai Tengah *Median*

Pada data berkelompok yang telah disusun dan diperoleh nilai distribusi frekuensinya maka kita dapat menghitung nilai median data tersebut. Perhitungan ini dapat dilakukan mengikut langkah-langkah berikut:

1. Menentukan menentukan letak median kelas interval dengan rumus sebagai berikut:

$$Me = \frac{n}{2}$$

2. Menghitung nilai median pada data berkelompok dengan memasukkan nilai media kelas interval pada langkah 1 di atas. Rumus menghitung median sebagai berikut:

$$M_e = Lm + \frac{w\left(\frac{n}{2} - cf\right)}{fm}$$

Dimana:

Me = Median

Lm = Batas bawah dari kelas interval dimana median berada (kelas median)

n = Banyaknya observasi

cf = Frekwensi kumulatif dari kelas interval sebelum kelas median

w = Lebar kelas interval dimana median berada

Contoh:

Tentukan median dari data kelompok dibawah ini Jawab:

Interval Kelas	f_i	
164,5 – 167,5	6	
167,5 – 170,5	7	→ *cf*
170,5 – 173,5	8	
173,5 – 176,5	11	→ *fm*
176,5 – 179,5	7	
179,5 – 182,5	6	
182,5 – 185,5	5	
Jumlah	50	

Jawaban

Menentukan kelas median $= \frac{n}{2} = \frac{50}{2} = 25$

Contoh:

Pada tabel disajikan sebuah sampel acak berupa laba bersih yang diraup oleh 60 rumah kecantikan SPA di Kota Denpasar. Laba Bersih dari 60 Rumah Kecantikan SPA di Kota Denpasar sebagai berikut:

Laba Bersih (Juta Rupiah)	Banyak SPA (unit)
5 - 9	4
10 - 14	7
15 - 19	13
20 - 24	22
25 - 29	8
30 - 34	6
Jumlah	60

Sumber: Data hipotetis

Berdasarkan data pada tabel di atas hitung nilai mediannya dan berikan interpretasi.

Jawaban

Tabel Perhitungan Median Laba Bersih per Bulan 60 Rumah Kecantikan SPA di Kota Denpasar.

Laba Bersih (Juta Rupiah)	Banyak SPA (f)	Tepi Kelas	F_c
		4.5	0
5 – 9	4		
		9.5	4
10 – 14	7		
		14.5	11
15 – 19	13		
		19.5	24
20 – 24	22		
		24.5	46
25 – 29	8		
		29.5	54
30 - 34	6		
		34.5	
Jumlah	60		

Dari tabel dapat diketahui: n = 60 , c = 5

$Me = \frac{n}{2} = \frac{60}{2} = 30$, yaitu terletak antara frekuensi komulatif 24 dan 46

Kelas mediannya adalah kelas ke-4, dengan kelas: 19.5 – 24.5

Jadi, L = 19.5, f_c = 24 dan f_m = 22.

Per rumus didapat,

$$M_e = Lm + \frac{w\left(\frac{n}{2} - cf\right)}{fm} = 19,5 + \frac{5(30-24)}{22}$$

$$M_e = 19,5 + 1,36$$

$$M_e = 20,86$$

Jadi, median dari laba bersih per bulan bagi 60 rumah kecantikan SPA di Kota Denpasar adalah Rp 20,86 juta. Nilai Mediannya = Rp 20,86 juta, memiliki arti bahwa 50% dari rumah kecantikan SPA (sampel) di Kota Denpasar, laba bersih per bulannya lebih kecil dari Rp 20,86 juta, dan 50% nya lagi lebih besar dari Rp 20,86 juta.

6.2.3 Modus Nilai Sering Muncul *Mode*

Pada kelas data berkelompok, kita akan lebih mudah memperoleh nilai modus pada jenis data ini. Hal ini karena pada data berkelompok kita hanya perlu melihat kelas mana yang mempunyai nilai distribusi frekuensi paling besar. Untuk memudahkan perhitungan nilai modus pada data berkelompok dapat menggunakan rumus berikut:

$$M_0 = Bb + w\left(\frac{d1}{d1 + d2}\right)$$

Contoh:

Tentukan nilai modus pada data tabel berikut:

Interval Kelas	f_i
164,5 – 167,5	6
167,5 – 170,5	7
170,5 – 173,5	8

Interval Kelas	f_i
173,5 – 176,5	11
176,5 – 179,5	7
179,5 – 182,5	6
182,5 – 185,5	5
Jumlah	50

Jawaban

$$M_0 = Bb + w\left(\frac{d1}{d1 + d2}\right)$$

$$M_0 = 176 + 7\frac{11 - 8}{(11 - 8) + (11 - 7)}$$

$$= 176 + 7\left(\frac{3}{7}\right)$$

$$= 176 + 3$$

$$= 179$$

Contoh:
Sebuah sampel nilai penjualan mingguan (juta rupiah) dari 60 penyalur barang antik di Kota Surabaya disajikan sebagai berikut:

Nilai Penjualan (Juta Rupiah)	Banyak Penyalur (f)
10 - 14	4
15 - 19	5
20 - 24	8
25 - 29	13
30 - 34	20
35 - 39	10
Jumlah	60

Berdasarkan data tersebut, hitunglah modusnya dan berikan interpretasi.

Jawaban

Tabel 4.19 Cara Perhitungan Modus Nilai Penjualan Mingguan 60 Penyalur Barang Antik

Nilai Penjualan (Juta Rupiah)	Banyaknya Penyalur (f)	Tepi Kelas
		9,5
10 - 14	4	
		14,5
15 - 19	5	
		19,5
20 - 24	8	
		24,5
25 - 29	13	
	) d_1	29,5
30 - 34	20 $\longleftarrow M_o$	
	) d_2	34,5
35 - 39	10	
		39,5
Total	60	

Berdasarkan tabel dapat diketahui bahwa frekuensi modusnya = 20, maka letak modus (L_{Mod}) pada kelas ke-5 (lihat tanda panah pada Tabel 4,19). Kelas nyatanya adalah 29,5 - 34,5.. Jadi. d_1 = 20 - 13 = 7, dan d_2 = 20 - 10 = 10 dan L = 29,5 serta c = 5.

Maka per rumus didapat,

$$M_0 = Bb + w\left(\frac{d1}{d1 + d2}\right)$$

$$M_0 = 29{,}5 + 5\left(\frac{7}{7 + 10}\right)$$

$$M_0 = 29{,}5 + 5(0{,}41)$$

$$M_0 = 29{,}5 + 2{,}05$$

Jadi, modus dari nilai penjualan mingguan 60 penyalur barang antik di Kota Surabaya adalah Rp 31,55 juta. Ini menunjukan bahwa nilai penjualan mingguan dari 60 penyalur barang antik tersebut yang paling banyak adalah nilai penjualan disekitar Rp 31,55 juta.

Contoh:

Dari sejumlah penderita typhus abdominalis yang dirawat di bangsal penyakit menular suatu Rumah Sakit, diperoleh data sebagai berikut:

Masa inkubasi (hari) dari 170 penderita *typhus abdominalis*

Masa inkubasi (hari)	**Jumlah penderita**
2	25
6	80
10	30
14	15
18	12
22	6
24*	2
	Jumlah = 170

* tidak ada pasien dengan masa inkubasi 30 hari atau lebih.

Hitung nilai rata-rata, median dan modus pada tabel masa inkubasi di atas.

Jawaban

Banyaknya pasien (f)	**Titik tengah** (x)	fx	$fx2$	**Frekuensi kumulatif** (cf)
25	4	100	400	25
80	8	640	5120	105
30	12	360	4320	135
15	16	240	3840	150
12	20	240	4800	162
6	24	144	3456	168
2	28	56	1568	170
Jumlah = 170		fx =1780	23504	

a. Rata-rata

$$\bar{x} = \frac{fx}{n} = \frac{1780}{170} = 10{,}47 \text{ hari}$$

b. Median

$$Md = Lm + \frac{\frac{n}{2} - cf}{fm} . w$$

$\frac{n}{2} = \frac{170}{2} = 85,$ kelas interval dimana median berada (kelas median) adalah: 6, maka

lm = 6

cf kelas interval sebelumnya = 25

fm = 80

w = 10 - 6 = 4

$$Md = 6 + \frac{\frac{170}{2} - 25}{80} . 4$$

$$Md = 6 + \frac{60}{80} . 4$$

$$Md = 6 + 3 = 9$$

c. Modus

Mo = 8, oleh karena frekuensi tertinggi dimiliki kelas interval 6 - dan titik tengah kelas interval ini adalah: 8.

6.2.4 Rata-rata Harmonis

Rata-rata harmonis data yang telah dikelompokkan dapat dihitung dengan rumus:

$$H = \frac{n}{\sum \frac{f_i}{m_i}}$$

H = rata-rata harmonis

n = banyaknya data

m_i = nilai tengah kelas yang ke-i

f_i = frekuensi kelas yang ke-i

6.3 UKURAN LETAK

6.3.1 Kuartil

Bila datanya telah dikelompokkan atau telah disusun dalam distribusi frekuensi atau tabel frekuensi, maka kuartil sekelompok data tersebut dapat dihitung dengan rumus sebagai berikut:

$$K_1 = B_{K_1} + \left(\frac{\frac{N}{4} - Cf_1}{f_{k_1}}\right).Ci$$

Yang menyatakan bahwa:

K_1 = Kuartil 1

N = Banyaknya data observasi (Σf)

Cf_1 = Brekuensi kumulatif kelas sebelum kelas kuartil 1

f_{K1} = Frekuensi kumulatif kelas kuartil 1

B_{K1} = Tepi kelas bawah kelas kuartil 1

Ci = Interval kelas

$$K_2 = B_{K_2} + \left(\frac{\frac{2N}{4} - Cf_2}{f_{k_2}}\right).Ci$$

Yang menyatakan bahwa:

K_2 = Kuartil 2

N = Banyaknya data observasi (Σf)

Cf_2 = Brekuensi kumulatif kelas sebelum kelas kuartil 2

f_{K2} = Frekuensi kumulatif kelas kuartil 2

B_{K2} = Tepi kelas bawah kelas kuartil 2

Ci = Interval kelas

K_2 = Nilainya sama dengan nilai median

$$K_3 = B_{K_3} + \left(\frac{\frac{3N}{4} - Cf_3}{f_{k_3}}\right).Ci$$

Yang menyatakan bahwa:

K_3 = Kuartil 3

N = Banyaknya data observasi (Σf)

Cf_3 = Brekuensi kumulatif kelas sebelum kelas kuartil 3

f_{K3} = Frekuensi kumulatif kelas kuartil 3

B_{K3} = Tepi kelas bawah kelas kuartil 3

Ci = Interval kelas

Contoh:

Omzet penjualan (juta rupiah) sampel acak 70 toko dalam sebuah komplek pertokoan di Kota Denpasar pada bulan lalu, disajikan sebagai berikut:

Tabel Omzet Penjualan 70 Toko dari Sebuah Komplek Pertokoan di Kota Denpasar

Omzet Penjualan (Juta Rupiah)	**Banyaknya Toko (Unit)**
20 - 29	1
30 - 39	4
40 - 49	7
50 - 59	13
60 - 69	25
70 - 79	15
80 - 89	5
Jumlah	70

Sumber: Data hipotetis

Berdasarkan data pada, hitung:

(a) Kuartil pertama (K_1), dan berikanlah interpretasi

(b) Kuartil ketiga (K_3), dan berikanlah interpretasi

Jawaban
Tabel Cara Menghitung K_1 dan K_3 dari Omzet Penjual an 70 Toko dalam Sebuah Komplek Pertokoan

Omzet Penjualan (Juta Rupiah)	Banyaknya Toko (Unit)	Tepi Kelas	f_c
		19,5	0
20 - 29	1	29,5	1
30 - 39	4	39,5	5
40 - 49	7	49,5	12
50 - 59	13	59,5	25
60 - 69	25	69,5	50
70 - 79	15	79,5	65
80 - 89	5	89,5	70
Jumlah	70		

Dengan x=1, dan n=70; Letak Q_1 yaitu antara frekuensi komulatif 12 dan 25 pada kelas ke-4 (kelas nyata) adalah 49,5 - 59,5.
Maka, Nilai Q_1 adalah sebagai berikut:

$$K_1 = B_{K_1} + \left(\frac{\frac{N}{4} - Cf_1}{f_{k_1}} \right) . Ci$$

$$K_1 = 49{,}5 + \left(\frac{\frac{70}{4} - 12}{13} \right) . 10$$

$$K_1 = 49{,}5 + 4{,}23$$

$$K_1 = 53{,}73$$

Jadi, kuartil pertama (Q_1) = Rp 53,73 juta. K_1 = Rp 53,73 juta, artinya kurang lebih 25% dari seluruh sampel toko tersebut omzet penjualannya kurang dari atau paling tinggi Rp 53,73 juta dan 75% nya lagi, omzet penjualan lebih dari atau paling rendah Rp 53,73 juta.

6.3.2 Desil

Desil data berkelompok dapat dihitung dengan rumus:

$$D_i = T_b + p\left(\frac{\frac{i}{10}n - F}{f}\right)$$

Dimana $\quad i \quad = 1,2,3,4,.....,9$

Dengan $\quad D_i$ = Desil ke-i

T_b = Tepi bawah interval kelas Di

p = Panjang kelas interval Di

n = Σf = Banyak data

F = Frekuensi kumulatif sebelum kelas Di

f = frekuensi pada kelas Di

Contoh:

Hitung nilai D_5 dan D_8 dari data berdistribusi kelompok berikut:

Interval	*F*	F_k
21 – 25	3	3
26 – 30	9	12
31 – 35	4	16
36 – 40	10	26
41 – 45	3	29
46 – 50	11	40

Jawaban

Desil ke-5 terletak pada $\left(\frac{5}{10}.40\right) = 20$ (kelas interval 36-40)

$$D_5 = 35{,}5 + \frac{5(20-16)}{10}$$

$$= 37{,}5$$

Desil ke-8 terletak pada $\left(\frac{8}{10}.40\right) = 32$ (kelas interval 46-50)

$$D_8 = 45{,}5 + \frac{5(32-29)}{11}$$
$$= 46{,}9$$

6.3.3 Persentil

Persentil data berkelompok dapat dihitung dengan rumus:

$$P_i = T_b + p\left(\frac{\frac{i}{10}n - F}{f}\right)$$

Dimana i = 1,2,3,4,.....,99
Dengan P_i = persentil ke-i
T_b = tepi bawah interval kelas Pi
p = panjang kelas interval Pi
n = Σf = banyak data
F = frekuensi kumulatif sebelum kelas Pi
f = frekuensi pada kelas Pi

Contoh:
Hitung nilai P_{25} dari data berdistribusi kelompok berikut:

Interval	*F*	F_k
21 – 25	3	3
26 – 30	9	12
31 – 35	4	16
36 – 40	10	26
41 – 45	3	29
46 - 50	11	40

Jawaban

Persentil ke-25 terletak pada $\left(\frac{25}{100}.40\right) = 10$ (kelas interval 26-30)

$$P_{25} = 25{,}5 + \frac{5(10-3)}{9}$$
$$= 29{,}4$$

6.4 UKURAN PENYEBARAN

6.4.1 Pengukuran Jarak (Range)

Range data yang telah dikelompokkan Bila datanya telah disusun dalam tabel frekuensi, rangenya dapat dihitung dengan rumus:

$$R = B_u - B_i$$

Dimana,

B_u = tepi kelas atas kelas yang terakhir

B_i = tepi kelas bawah kelas yang pertama

Range data observasi berkelompok (*grouped data*) adalah data selisih antara tepi kelas atas kelas yang terakhir dengan tepi kelas bawah kelas pertama.

Contoh:

Berikut ini adalah data mengenai nilai 30 peserta ujian Matematika di SMA Mercu Buana Yogyakarta

TABEL 2.1 ***Nilai Peserta Ujian Matematika***

Nilai	FREKUENSI (f)
40	6
50	10
60	4
70	4
80	2
90	4

Range nilai 30 peserta ujian matematika dapat ditentukan dengan menggunakan Rumus:

$$R = B_u - B_i$$

Dengan nilai-nilai

B_u = 99,5 (tepi kelas atas kelas yang terakhir)

B_i = 39,5 (tepi kelas bawah kelas yang pertama)

Sehingga besarnya Range (R)

$$R = B_u - B_i$$

$$R = 99{,}5 - 39{,}5 = 60$$

6.4.2 Pengukuran Deviasi Kuartil

Rumus penghitungan deviasi kuartil pada data acak adalah sama dengan rumus pada perhitungan data berkelompok. Sehingga, rumus yang digunakan untuk menghitung deviasi kuartil adalah sebagai berikut:

$$dQ = \frac{Q_3 - Q_1}{2}$$

6.4.3 Pengukuran Deviasi Rata-rata (*Mean Deviation*)

Apabila nilai-nilai observasi pada data berkelompok sudah dikelompokkan ke dalam bentuk distribusi frekuensi, maka deviasi rata-ratanya dapat dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum_{i=1}^{k} f_i |m_i - \bar{x}|}{n}$$

Dimana:

m_i = Titik tengah kelas frekuensi

f_i = Frekuensi dari kelas distribusi ke-i

k = Jumlah kelas distribusi

Dalam beberapa kondisi tertentu, median dapat digunakan sebagai pengukuran rata-rata secara memuaskan. Deviasi rata-rata sebuah distribusi dapat juga diukur dari median distribusi yang bersangkutan seperti dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum|x_i - md|}{n}$$

Atau;

$$d_{\bar{x}} = \frac{\sum f_i|m_i - md|}{n}$$

Umumnya deviasi rata-rata merupakan pengukuran dispersi yang lebih baik jika dibandingkan dengan jarak atau deviasi kuartil. Hasil pengukuran deviasi rata-rata mencerminkan dispersi tiap-tiap nilai observasi dari rata-ratanya dan bukan hanya tergantung pada kedua nilai ekstrim.

Contoh:
Dari data tunggal dibawah ini, rubahlah menjadi data kelompok:

40	50	70	55
55	72	66	60
60	54	85	65
45	67	80	75
70	80	55	80

Dan carilah deviasi rata-ratanya.

Jawaban
Data setelah dikelompokkan

Nilai	F	Mi
4047	2	43,5
4855	5	51,5
5663	2	59,5
6471	5	67,5
7279	2	75,5
8087	4	83,5
Σ	20	

$$Median\ (Md) = M_0 + P\left(\frac{\frac{n}{2} - fm}{fc}\right)$$

$$Median\ (Md) = 63{,}5 + 8\left(\frac{10 - 9}{5}\right)$$

$$Median\ (Md) = 63{,}5 + 8(0{,}2)$$

$$Median\ (Md) = 65{,}1$$

$$d_{\bar{x}} = \frac{\sum f_i\ |x_i - md|}{n}$$

$$d_{\bar{x}} = \frac{|(43{,}5 - 65{,}1).\,2 + (51{,}5 - 65{,}1).\,5 + \cdots + (83{,}5 - 65{,}1).\,4|}{20}$$

$$d_{\bar{x}} = \frac{228{,}8}{20} = 11{,}44$$

Contoh:

Berdasarkan data pada Tabel di bawah, hitunglah deviasi rata-rata omzet penjualan 70 toko tersebut.

Tabel Perhitungan Deviasi Rata-rata Omzet Penjualan 70 Toko

Omzet Penjualan (Juta Rp)	f_i	m_i	$m_i - \bar{x}$	$f_i\lvert m_i - \bar{x}\rvert$	$f_i \cdot m_i$
20 - 29	1	24.5	-37.43	37.43	24.5
30 - 39	4	34.5	-27.43	109.2	138.0
40 - 49	7	44.5	7.43	122.01	311.5
50 - 59	13	54.5	-7.43	96.59	708.5
60 - 69	25	64.5	2.57	64.25	1612.5
70 - 79	15	74.5	12.57	188.55	1117.5
80 - 89	5	84.5	22.57	112.85	422.5
Jumlah	70			731.4	4,335

Jawaban

Dari Tabel di atas dapat diketahui bahwa n = $\sum f_i$ = 70, $\sum f_i\, m_i$= 4,335

Dihitung terlebih dahulu dengan rumus sebagai berikut:

$$d_{\bar{x}} = \frac{\sum_{i=1}^{k} f_i m_i}{n}$$

$$d_{\bar{x}} = \frac{4335}{70}$$

$$d_{\bar{x}} = 61{,}93$$

Dari dapat diketahui bahwa $\sum f_i|m_i - \bar{x}|$ = 731.4

Selanjutnya per rumus deviasi rata-rata dihitung dan didapat hasil sebagai berikut:

$$d_{\bar{x}} = \frac{\sum f_i|m_i - \bar{x}|}{n}$$

$$d_{\bar{x}} = \frac{731{,}4}{70}$$

$$d_{\bar{x}} = 10.45$$

Jadi, deviasi rata-rata omzet penjualan bagi 70 toko (sampel) tersebut adalah Rp 10.45 juta.

Interpretasi nilai deviasi rata-rata = Rp 10.45 juta memiliki arti bahwa secara rata-rata ketujuh puluh omzet penjualan toko tersebut menyimpang sebesar Rp10.45 juta dari nilai rata-ratanya (= Rp 61.93 juta.)

6.5 PENGUKURAN VARIANS DAN DEVIASI STANDAR

6.5.1 Variansi dan Deviasi Standar Data Berkelompok

Varians dan deviasi standar dari data yang telah dikelompokkan dapat dihitung dengan rumus sebagai berikut"

- **Varians** dari data sampel yang telah dikelompokkan:

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

- **Deviasi standar** dari data sampel yang telah dikelompokkan:

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

Dimana:

x_i = Titik tengah tiap-tiap kelas

f_i = Jumlah frekuensi kelas

6.5.2 Variansi dan Deviasi Standar dengan Cara Transformasi

Seperti halnya dengan mencari nilai mean data kelompok. Kita juga dapat mencari nilai variansi dapat dicari dengan cara transformasi.

$$u_i = x_i - a$$

Dimana:

x_i : Titik tengah interval kelas ke-i

α : Sembarang harga titik tengah interval kelas (**biasanya yang memiliki frekuensi terbanyak**)

sehingga rumus **VARIANSI ()** adalah:

$$s^2 = c^2 u^2$$

c = Lebar kelas/panjang kelas

Dimana:

$$s_{\bar{u}}^2 = \frac{1}{n-1}\sum_{i=1}^{k} f_i (u_i - \bar{u})^2$$

Atau dapat juga ditulis:

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} f_i u_i^2 - \frac{1}{n}\left(\sum_{i=1}^{k} f_i u_i\right)^2\right]$$

Contoh:

Dari data tinggi badan (cm) 50 mahasiswa Pendidikan Matematika FKIP Universitas Mercu Buana Yogyakarta didapat data:

TABEL 1. ***Perhitungan variansi data berkelompok***

Interval kelas	x_i	u_i	f_i	u_i^2	$f_i u_i$	$f_i u_i^2$
164,5 – 167,5	166	166-175=-9	6	81	6*-9=-54	6*81 =486
167,5 – 170,5	169	169-175=-6	7	36	7*-6=-42	7*36 = 252
170,5 – 173,5	172	-3	8	9	-24	72
173,5 – 176,5	175	0	11	0	0	0
176,5 – 179,5	178	178-175= 3	7	9	21	63
179,5 – 182,5	181	6	6	36	36	216
182,5 – 185,5	184	9	5	81	45	405
Jumlah			50		**-18**	**1494**

Berdasarkan tabel 1 dengan menggunakan rumus transormasi, maka variansinya:

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} f_i u_i^2 - \frac{1}{n}\left(\sum_{i=1}^{k} f_i u_i\right)^2\right]$$

$$s = \frac{1}{50-1}\left(1494 - \frac{1}{50}(-18)^2\right) = 30,35$$

$$s = \sqrt{30,4} = 5,50$$

Bab 7
Analisis Data Berkala

7.1 PENGERTIAN ANALISIS DERET BERKALA

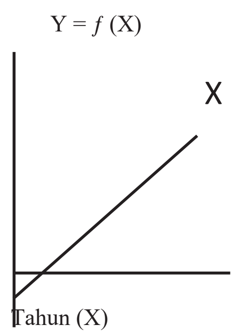
Data deret berkala adalah data yang dikumpulkan dari satu periode ke periode tertentu untuk menggambarkan perkembangan suatu kegiatan seperti perkembangan fluktuasi harga, perkembangan produksi, konsumsi, jumlah penduduk, angka kemiskinan, luas panen padi, harga saham, Curah hujan dan lain-lain. Analisis data berkala (trend) atau *Time Series Analysis* merupakan teknik statistik yang digunakan untuk memprediksi masa mendatang untuk data yang sudah tersedia. Kegunaan dari analisis data berkala adalah sebagai alat analisis *forcasting* atau peramalan yang dapat melihat kondisi masa mendatang. analisis deret berkala merupakan analisa yang menggunakan metode kuantitatif untuk mempelajari pola data dimasa lampau untuk memprediksi dan menganalisa dari data masa lampau tersebut, untuk memperolah informasi atau suatu gambaran peubah, dan hasilnya bertujuan untuk membuat suatu rencana dimasa yang akan datang.

7.2 KOMPONEN DERET BERKALA

Peubah deret berkala dipengaruhi oleh 4 (empat) komponen deret berkala. Menurut Levin (1981) dan Black (2011) adalah Tren sekuler (T), Variasi musim (S), Variasi siklis (C), dan Variasi residu (I).

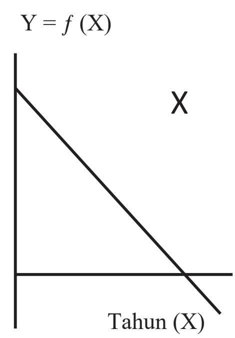
7.2.1 Trend Sekuler (*Secular Trends)*

Trend atau gerakan jangka panjang adalah suatu gerakan yang menunjukan suatu perkembangan yang dilihat dari suatu arah yang menunjukan kenaikan atau penurunan pada jangka panjang, yang didapatkan dari rata-rata perubahan dari satu periode ke periode tertentu, dan jangka waktu yang biasa digunakan adalah 10 tahun keatas. Trend arah naik maka akan menunjukkan trend positif, sedangkan trend menurun akan menunjukkan trend negatif.



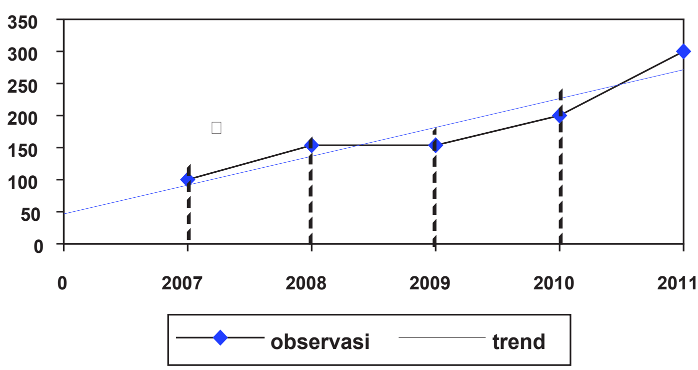
Trend Positif

GAMBAR 7.1



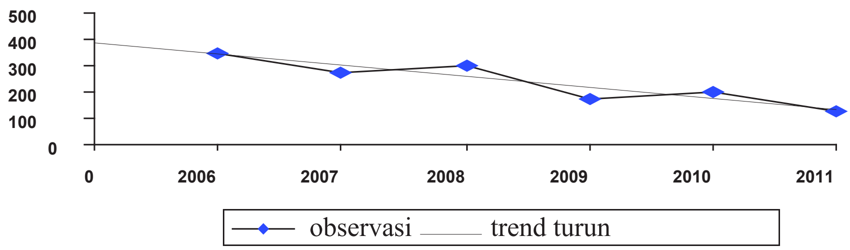
Trend Negatif

GAMBAR 7.2



GAMBAR 7.3 ***Trend Naik Biaya Hidup Kurun Waktu Tahun 2007- 2011***

Biaya hidup seseorang akan cenderung selalu meningkat setiap tahunnya. hal ini disebabkan karena semakin banyaknya kebutuhan dan disertai dengan adanya kenaikan harga. Kondisi ini menyebabkan trend mengalami kenaikan.



GAMBAR 7.4 ***Trend Turun Dari Luas Lahan Pertanian Pada Daerah "X" Kurun waktu 2006 - 2011***

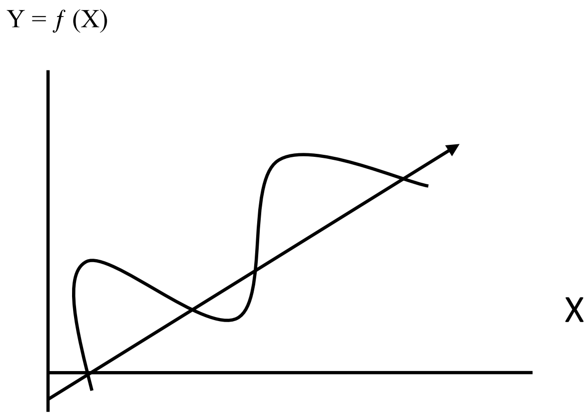
Setiap tahun lahan pertanian di perkotaan maupun dipedesaan semakin sempit. Lahan yang semakin sempit ini disebabkan karena adanya peralihan fungsi lahan. Peralihan fungsi lahan pertanian ini digunakan untuk bangunan real estat, kawasan industri, gedung-gedung bertingkat, sehingga semakin sedikit lahan yang dijadikan sebagai lahan pertanian, kondisi ini menyebabkan trend mengalami penurunan.

7.2.2 Gerakan/Variasi Siklis (*Cyclical Variations*)

Variasi siklis merupakan suatu variasi yang dipengaruhi oleh fluktuasi ekonomi jangka panjang. Variasi/gerakan siklis ini terjadi berulang setlah jangka waktu tertentu. Variasi siklis memiliki gerakan naik atau turun secara periodik pada jangka waktu yang panjang, yaitu 3 tahun, 5 tahun, 10 tahun, 20 tahun, 25 tahun dan bisa melebihi. Namun, bisa juga tidak terjadi kejadian yang berulang saat jangka waktu sama. Suatu kegiatan ekonomi yang menunjukkan variasi siklis adalah industri konstruksi bangunan yang mempunyai gerakan siklis 15-20 tahun. Contoh lain dari gerakan siklis adalah *Business Cyclis*.

Dalam *Business Cyclis* dibagi dalam 4 periode yaitu:

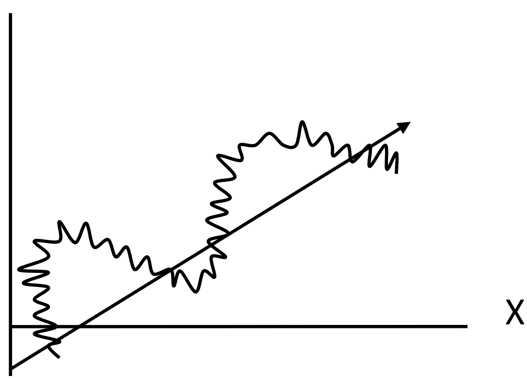
1. Masa Pemulihan (*Revival Phase*)
2. Masa Kemakmuran (*Prosperity Phase*)
3. Masa Kemuduran/krisis (*Crisis Phase*)
4. Masa Kehancuran (*Depression Phase*)



GAMBAR 7.5 ***Trend Jangka Panjang dan Gerakan/Variasi Siklis***

7.2.3 Gerakan/Variasi Musim (*Seasonal Variations*)

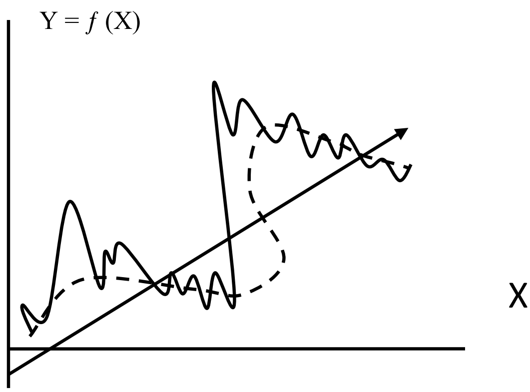
Suatu gerakan yang menunjukan perkembangan fluktuasi secara teratur meningkat atau turun pada waktu musiman atau bulan tertentu dalam jangka waktu kurang dari setahun. Biasanya terjadinya perkembangan fluktuasi ini pada saat musim lebaran hari raya, pergantian tahun, melimpahnya hasil produksi pada komoditas tertentu pada musim tertentu. Sebagai contoh produksi pertanian pada komoditas buah mangga atau buah durian di pasar atau dimall pada masa musim panen atau dibulan-bulan tertentu. Peningkatan penjualan baju dimusim lebaran hari raya. Dan penjualan alat-alat tulis dan tas sekolah saat awal dimulainya semester baru disekolah.



GAMBAR 7.6 ***Trend Jangka Panjang, Gerakan Siklis, dan Musiman***

7.2.4 Gerakan/Variasi Residu (*Irregular Vaiations*)

Variasi residu adalah suatu variasi yang tidak teratur dan sulit untuk diramalkan dan diprediksi. Variasi ini bisa disebabkan karena faktor kebetulan pada suatu kegiatan. Variasi ini memiliki sifat sporadis, misalnya terjadi siklus perekonomian yang naik turun yang disebabkan oleh bencana alam seperti banjir, gunung meletus, banjir dan longsor, peperangan, kelaparan, gejolak politik yang sedang tidak stabil dan yang lainnya. Fenomena ini dapat mempengaruhi laju perekonomian seperti perkembangan produksi, investasi, perdagangan.



GAMBAR 7.7 ***Trend Jangka Panjang, Gerakan Siklis, Musiman, dan Residu/acak***

Analisis data berkala umumnya terdiri dari uraian secara matematis tentang komponen-komponen yang menyebabkan gerakan-gerakan atau variasi-variasi yang tercermin dalam fluktuasi. Apabila gerakan Trend jangka panjang, Variasi Siklis, Variasi Musiman dan Residu masing-masing diberikan symbol *T, C, S,* dan *I*, maka data berkala *Y* merupakan hasil kali dari 4 komponen tersebut, yaitu:

$$Y = T \times C \times S \times I \qquad (7.1)$$

Selain itu, ahli statistic juga menganggap bahwa data berkala merupakan hasil penjumlahan dari 4 komponen tersebut, yaitu:

$$Y = T + C + S + I \qquad (7.2)$$

7.3 TREND LINEAR

7.3.1 Persamaan Tren Linear

Trend linear memiliki persamaan sebagai berikut:

$$Y' = a + bX \tag{7.3}$$

$$Y' = a - bX \tag{7.4}$$

Keterangan: Y = peubah terikat
X = peubah bebas (dalam hal ini, x = waktu)
a = intersep Y, merupakan bilangan konstan
b = slope/arah garis tren

7.3.2 Metode Trend Linear

Terdapat beberapa metode yang digunakan untuk menyusun atau menggambarkan garis trend linear, yaitu:

1. Metode bebas (*Freehand Method*)
2. Metode rata-rata semi (*Semi Average Method*)
3. Metode rata-rata bergerak (*Moving Average Method*)
4. Metode kuadrat terkecil (*Least Squares Method*)

1. Metode bebas (*Freehand Method*)

Metode bebas adalah metode yang paling mudah, tetapi sifatnya sangat subjektif. Metode ini menerapkan garis trend tanpa menggunakan rumus matematika. Untuk menentukan gerak tren dengan metode ini, langkah-langkah yang harus dilakukan adalah membuat sumbu tegak Y dan sumbu mendatar X. selanjutnya, buat *scatter diagram*. Data hasil pengamatan atau observasi dibuat diagram pancar dan grafik, setelah itu, tariklah garis lurus mendekati semua titik koordinat pada diagram pancar secara bebas. Walaupun merupakan metode yang paling mudah dan tanpa menggunakan rumus matematika, penggambaran garis trend dengan metode ini dianggap kurang memenuhi syarat ilmiah dan jarang digunakan.

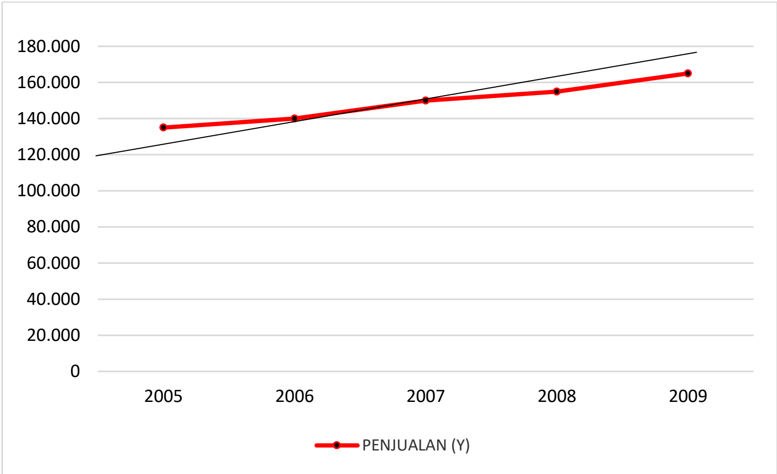
Contoh:

Diketahui Sebuah Perusahaan PT. SIDOMUKTI membuat sebuah peramalan penjualan untuk tahun 2005-2009 dengan menggambarkan garis trend. Dibawah ini terdapat data penjualannya dan terdapat grafik yang tersaji:

TABEL 7.1 *Data Penjualan PT. Sidomukti Tahun 2005-2009*

TAHUN (X)	PENJUALAN (Y)
2005	135.000
2006	140.000
2007	150.000
2008	155.000
2009	165.000
2010	?

Selanjutnya melakukan peramalan penjualan tahun 2010 dengan menggunakan metode peramalan Trend Bebas. Langkah pertama yang harus dilakukan adalah dengan memplotkan seluruh data ke grafik sumbu kartesius dimana sumbu vertical mewakili penjualan dan sumbu horisontak mewakili tahun. Setelah itu, dibuatlah garis trend bebas paling tidak mendekati semua titik koordinat yang sebagaimana terlihat pada gambar dibawah ini.



GAMBAR 7.8 *Grafik Trend Bebas*

Dari metode trend bebas diatas, bahwa peramalan penjualan untuk tahun 2010 adalah 175.000 unit.

2. Metode Setengah rata-rata (*Semi Average Method*)

Cara menentukkan tren metode setengah rata-rata adalah:

1. Bagilah data deret waktu tersebut menjadi dua kelompok, yaitu kelompok 1 dan kelompok 2 yang memiliki jumlah data yang sama. Jika jumlah tahun nya genap, langsung dibagi menjadi dua. Hasilnya ada dua kemungkinan, pertama dua kelompok data genap, dan kelompok data ganjil. Bila jumlah tahunnya ganjil, tahun pertengahan dihilangkan saja atau dimasukkan kedalam kedua kelompok. Hasilnya ada dua kemungkinan, pertama dua kelompok data genap, dan kelompok data ganjil.
2. Carilah rata-rata tiap kelompok (X_1 dan X_2) rata-rata hitung ini, disebut setengah rata-rata, dan letaknnya pada tahun (waktu) pertengahan tiap kelompok.
3. Nilai setengah rata-rata pada masing-masing kelompok dapat dianggap sebagai nilai tren per 30 Juni masing-masing periode dasar (periode dasar = tahun yang memuat nilai setengah rata-rata).
4. Menghitung perubahan trend dengan rumus:

$$b = \frac{x2 - x1}{n} \tag{7.5}$$

Catatan:

b = rata-rata perubahan peubah y per satuan waktu (tahun) jika b > 0 = pertambahan dan bila

b < 0 = penurunan.

n = banyaknya unit tahun antara tahun dasar (tahun dasar 1 sampai dengan tahun dasar 2.

5. Persamaan garis trennya adalah:

$$Y' = a + bX \tag{7.6}$$

Catatan:

Y' = nilai tren periode tertentu

a = nilai tren periode dasar

b = rata-rata perubahan nilai tren per satuan waktu (tahun)

Nilai tren pada tahun dasar (x=0), otomatis sama dengan nilai rata-rata tiap kelompok.

Contoh:
Berikut ini data omzet penjualan (miliar rupiah) PT. INDIRAJAYA untuk tahun 2010-2015. data tersebut diminta untuk membuat peramalan omzet penjualan untuk tahun 2016 dan 2017 dengan menggunakan metode *semi average*.

TABEL 7.2 ***Data Omzet Penjualan (Miliar Rupiah) PT. INDIRAJAYA Tahun 2010-2015***

Tahun	Penjualan (Ribu Ton)
2010	200
2011	250
2012	300
2013	450
2014	550
2015	650

Tentukan:
a) Dengan metode semi rata-rata, tentukan persamaan trennya.
b) Berikan interpretasi nilai b yang diperoleh.
c) Tentukan nilai tren untuk masing-masing tahun.
d) Perkirakan atau ramalkan omzet penjualan PT. INDIRAJAYA pada tahun 2016 dan 2017.
e) Buatlah grafiknya.

Jawaban

a.) Menentukan persamaan tren:
n = 6 (genap) = n/2= 6/2 = 3 (dua kelompok data ganjil yang masing-masing terdiri atas 3 buah data.
Tabel Perhitungan tren omzet penjualan PT. INDIRAJAYA dengan Metode Semi Rata-rata tahun 2010-2015

Tahun	Jumlah Omzet Penjualan (Ribu Ton)	Rata-rata per kelompok (1000 ton)	Nilai Tren (Ribu ton)
2010 2011 2012	200 250 300 $\Big\}\ x_1$	$x_1 = \frac{750}{3} = 250$	250
2013 2014 2015	450 550 650 $\Big\}\ x_2$	$x_2 = \frac{1650}{3} = 550$	550

Dari tabel diatas, dapat diketahui bahwa:

a = 250 (Nilai tren 2011)

a = 550 (Nilai tren 2014)

n = 2014 – 2011 = 3

Nilai b dihitung dan didapatkan:

$$b = \frac{x2 - x1}{n} = \frac{550 - 250}{3} = 100$$

Jadi Nilai persamaan trennya adalah:

1) Y' = a + bX
   = 250 + 100X (tahun dasar 2011)
2) Y' = a + bX
   = 550 + 100X (tahun dasar 2014)

b.) Interpretasi:
Nilai b = 100, yang artinya bahwa rata-rata kenaikan omzet penjualan PT. INDIRAJAYA tersebut sebesar Rp 100 miliar per tahun selama periode tahun 2010-2015.

c.) Menentukan nilai tren masing-masing tahun:
Nilai tren dapat ditentukan berdasarkan salah satu persamaan tren pada butir (a). Nilai tren pada masing-masing tahun akan dihitung berdasarkan persamaan tren dengan tahun dasar 2011. Pertama, dibuat skala nya X = 0 terlebih dahulu dan diletakan pada tahun dasar (tahun 2011). Skala x sebagai berikut:

Tahun	2010	2011	2012	2013	2014	2015	2016	2017
X (1 Thn)	-1	0	1	2	3	4	5	6

Berdasarkan persamaan tren pada butir (a) yaitu Y' = 250 + 100X, maka nilai tren masing-masing tahun dapat dihitung sebagai berikut:

Persamaan tren: Y' = 250 + 100X (tahun dasar 2011)

Tahun 2010: Y' = 250 + 100(-1)
= 150

Tahun 2011: Y' = 250 + 100(0)
= 250

Tahun 2012: Y' = 250 + 100(1)
= 350

Tahun 2013: Y' = 250 + 100(2)
= 450

Tahun 2014: Y' = 250 + 100(3)
= 550

Tahun 2015: Y' = 250 + 100(4)
= 650

Tabel Jumlah Omzet Penjualan dan Nilai Tren PT. INDIRAJAYA tahun 2010-2015

Tahun	Jumlah Omzet Penjualan (Miliar Rp)	Nilai Tren
2010	200	150
2011	250	250
2012	300	350
2013	450	450
2014	550	550
2015	650	650

d.) Untuk memperkirakan atau meramalkan jumlah omzet penjualan PT. INDIRAJAYA pada tahun 2016 dan 2017 Sebagai berikut:

Persamaan tren: Y' = 250 + 100X (tahun dasar 2011)

Tahun 2016: Y' = 250 + 100(5)

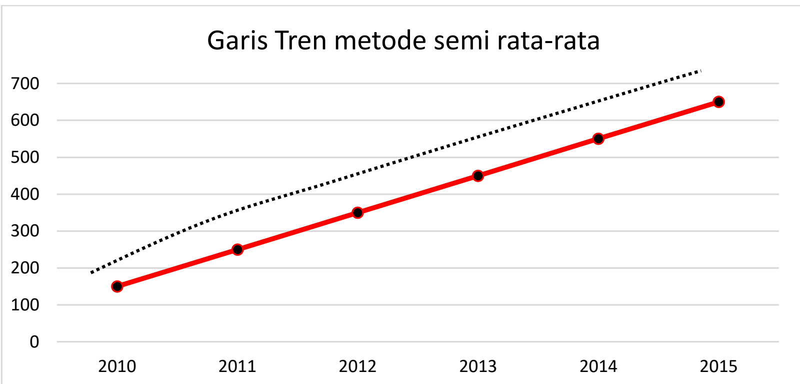
= 750

Tahun 2017: Y' = 250 + 100(6)

= 850

Berdasarkan perhitungan peramalan diatas, bahwa jumlah omzet penjualan PT. INDIRAJAYA pada tahun 2016 dan 2017 diperkirakan sebesar 750 dan 850 Milliar Rupiah.

e.) Gambar 7.9 Grafik Nilai Tren



3. Metode rata-rata bergerak (*Moving Average Method*)

Metode rata-rata bergerak adalah suatu metode peramalan untuk menghitung rata-rata nilai runtun dan digunakan untuk memperkirakan nilai pada tahun atau periode berikutnya. Metode ini ini menggunakan penjumlahan dan mencari nilai rata-rata dari sejumlah periode tertentu. Dengan menggunakan metode ini, pengaruh factor musim dan factor lainnya dapat hilangkan sehingga tren dapat dihitung. Nilai data tahunan seperti triwulan, bulanan, kwartal atau semesteran akan diganti dengan nilai rata-ratanya. Metode ini lebih baik digunakan untuk data yang bersifat stabil. Metode ini terdiri dari dua pola, yaitu: pola gerak ganjil (tarif N ganjil): 3, 5, 7 dan Pola Gerak Genap (taraf N genap): 2, 6, 8.

Rumus *Moving Average* (Metode rata-rata bergerak)

$$MA = \sum X \,/\, \text{Jumlah Periode} \qquad (7.7)$$

Keterangan:

MA	=	*Moving Average*
$\sum X$	=	Keseluruhan Penjumlahan dari semua data periode waktu yang diperhitungkan
Jumlah Periode	=	Jumlah Periode rata-rata bergerak

Atau bisa ditulis dengan rumus:

$$MA = (n1 + n2 + n3 + \ldots)/\, n \qquad (7.8)$$

Keterangan:

MA = *Moving Average*
n1 = data periode pertama
n2 = data periode kedua
n3 = data periode ketiga
n = Jumlah periode rata-rata bergerak

4. Metode kuadrat terkecil (*Least Squares Method*)

Metode kuadrat terkecil adalah metode yang menentukan garis trend yang mempunyai jumlah terkecil kuadrat selisih data asli dengan data garis tren. Cara kerja metode terkecil adalah meminimumkan jumlah kuadrat penyimpangan (selisih) nilai variable bebasnya (Y_i) dengan nilai tren/nilai ramalannya (Y_i) yaitu $\sum(_{Yt - Yt)}{}^2 = \sum e^2$. dengan adanya bantuan kalkulus yaiitu derivasi parsial $\sum_{(Yt - Yt)}{}^2 = \sum e^2$ diminimumkan, akan diperoleh dua buah persamaan normal secara simultan, sebagai berikut:

$$\sum Y_i = na + b\sum X_i \qquad (7.9)$$

$$\sum X_i = a\sum X_i + b\sum X_i^2 \qquad (7.10)$$

Dengan menyelesaikan kedua persamaan normal ini, nilai a dan b dari persamaan tren Y' = a + bX dapat dihitung. Agar perhitungan menjadi lebih sederhana pemberian kode pada nilai x (tahun) diupayakan sehingga $\sum X_i = 0$. Persamaan normal disederhanakan lagi:

a. Cara Cepat ($\sum X \neq 0$):
   Harus ada koding, $X_1 = 0$ [ada koding pertama, $X_2 = 1$ dam seterusnya.

$$a = \frac{\sum X^2 \sum Y - \sum X \sum XY}{n\sum X^2 - (\sum X)^2} \tag{7.11}$$

$$b = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2} \tag{7.12}$$

b. Cara Pendek ($\sum X = 0$):
   X untuk N ganjil = ….,-2,-1,0,1,2,….
   X untuk N genap = …..,-2,5;-1,5;-0,5;0,5;1,5;2,5;…..

$$a = \frac{\sum Y}{n} \tag{7.13}$$

$$b = \frac{\sum X\,Y}{\sum X2} \tag{7.14}$$

Setelah memasukan nilai a dan b, selanjutnya dengan memasukan nilai a dan b ke dalam persamaan, maka persamaan tren linear dapat disusun dibawah ini:

$$Y' = a + bX \tag{7.15}$$

Y’ = nilai taksiran atau nilai tren
a = intersep, yaitu besarnya nilai Y jika nilai X = 0
b = slope garis tren, yaitu perubahan variable Y untuk setiap perubahan satu unit peubah X.
X = periode waktu

Contoh:
Dengan metode Moving Average dan Least Squares Method, pada PT. CASA bergerak dibidang manufaktur makanan produksi Susu (ton) dan pada tahun 2014-2020 Produksinya sebagai berikut:

TABEL 7.3 ***Produksi Susu PT. CASA Tahun 2014-2020***

Tahun	Y
2014	7
2015	10
2016	16
2017	13
2018	22
2019	25
2020	31

Tentukan:

a.) Nilai tren produksi susu dengan menggunakan *Moving Average Method?*

b.) Persamaan trend menggunakan *Least Square Method*?

Jawaban

n = 7

Y = 7, 10,16,13,22,25,31

a.) *Moving Average Method* dan rata-rata bergerak 3 tahun

Tahun	Y	Jumlah bergerak selama 3 tahun	Rata-rata bergerak selama 3 tahun
2014	7	-	-
2015	10	7+10+16 = 33	33/3 = 11
2016	16	10+16+13 = 39	39/3 = 13
2017	13	16+13+22 = 51	51/3 = 17
2018	22	13+22+25 = 60	60/3 = 20
2019	25	22+25+31 = 78	78/3 = 26
2020	31	-	-

Kesimpulannya, untuk nilai trend dengan menggunakan *Moving Average Method* untuk produksi susu dari tahun 2014-2020 berturut-turut adalah 11, 13, 17, 20, 26 per ton.

b.) *Least Squares Method*

		Cara Panjang			Cara Pendek		
Tahun	**Y**	**X**	**X^2**	**XY**	**X**	**X^2**	**XY**
2014	7	0	0	0	-3	9	-21
2015	10	1	1	10	-2	4	-20
2016	16	2	4	32	-1	1	-16
2017	13	3	9	39	0	0	0
2018	22	4	16	88	1	1	22
2019	25	5	25	125	2	4	50
2020	31	6	36	186	3	9	93
Jumlah	124	21	91	480	0	28	108

LSM Cara Panjang =

$$a = \frac{\sum X2 \sum Y - \sum X \sum XY}{n \sum X2 - (\sum X)2} = \frac{91(124) - 21(480)}{7(91) - 21^2} = \frac{1{,}204}{196} = 6{,}1428571429$$

$$b = \frac{n \sum XY - \sum X \sum Y}{n \sum X2 - (\sum X)2} = \frac{7(480) - 21(124)}{7(91) - 21^2} = \frac{756}{196} = 3{,}8571428571$$

LSM Cara Pendek =

$$a = \frac{\sum Y}{n} = \frac{124}{7} = 17{,}7142857143$$

$$b = \frac{\sum X\,Y}{\sum X2} = \frac{108}{28} = 3{,}8571428571$$

Bab 8
Angka Indeks

8.1 PENGANTAR

Fokus pada Bab ini adalah untuk memberikan gambaran kepada pembaca tentang pentingnya angka indeks dalam proses penelitian. Data yang telah diperoleh, diolah dan disajikan dengan benar dapat dianalisis. Kegunaan angka indeks dalam analisis adalah untuk menyajikan data dalam bentuk yang sama. Sehingga, angka indeks dapat memberikan gambaran perubahan atau perkembangan terhadap data yang diperoleh. Pokok bahasan dalam bab ini meliputi 1) Pengertian angka indeks, (2) Jenis-jenis angka indeks, (3) Penyusunan Angka Indeks, (4) Metode perhitungan angka indeks, (5) Angka indeks tidak tertimbang, (6) Angka indeks tertimbang, (7) Angka indeks rantai, dan (8) Angka indeks untuk deflasi.

8.2 PENGERTIAN ANGKA INDEKS

Angka indeks adalah suatu angka yang menggambarkan perubahan relatif suatu kegiatan dalam dua waktu yang berbeda. Angka indeks dapat digunakan untuk mengukur perubahan secara kuantitatif dalam dua waktu yang berlainan. Menghitung angka indeks diperlukan dua macam waktu yaitu: waktu dasar (base period) dan waktu yang bersangkutan (current period). Waktu dasar adalah waktu yang digunakan sebagai dasar perbandingan, sedangkan waktu yang bersangkutan adalah waktu yang dipergunakan sebagai perbandingan terhadap kegiatan pada waktu dasar. Pemilihan waktu dasar biasanya memperhatikan kondisi perekonomian yng normal dan tidak terlalu jauh dengan tahun yang dibandingkan. Angka indeks umumnya digunakan untuk menghitung Indeks Harga, indeks kuantitas, dan indeks nilai.

8.3 JENIS-JENIS ANGKA INDEKS

Dalam bidang ekonomi pada dasarnya ada 3 (tiga) jenis angka indeks, yaitu (1) indeks harga, (2) indeks kuantitas, dan (3) indeks nilai

8.3.1 Indeks Harga (*Price Indeks*)

Indeks harga adalah angka yang dapat dipakai untuk melihat perubahan mengenai harga-harga barang, baik harga sejenis barang maupun sekelompok barang dalam waktu dan tempat yang sama ataupun berlainan.

8.3.2 Indeks Kuantitas (*Quantity Indeks*)

Indeks kuantitas adalah angka yang dapat dipakai untuk melihat perubahan mengenai kuantitas sejenis barang atau sekelompok barang yang dihasilkan (diproduksi), dijual, dikonsumsi, diekspor, dan sebagainya dalam waktu yang sama atau berlainan.

8.3.3 Indeks Nilai *(Value Indeks)*

Indeks nilai adalah angka yang dapat dipakai untuk melihat perubahan nilai uang dari suatu barang yang diproduksi, diekspor, diimpor, dikonsumsi dan sebagainya dalam waktu dan tempat yang sama atau berlainan. Nilai ini dapat diperoleh dari hasil perkalian antara harga per unit barang dengan kuantitasnya. Sebagai contoh, misalnya indeks **biaya hidup** pada dasarnya merupakan nilai pengeluaran konsumsi setiap keluarga, yang tak lain dari hasil perkalian antara harga per unit dan kuantitas barang yang dikonsumsi. Demikian juga halnya dengan **nilai produksi,** yang tak lain merupakan hasil perkalian antara harga per unit dengan kuantitas barang yang diproduksi.

Pengertian angka indeks yang lainnya dalam bidang ekonomi merupakan kombinasi diantara dua, dari tiga angka indeks tersebut. Selanjutnya dalam buku ini akan dibahas lebih banyak mengenai **indeks harga,** oleh karena indeks inilah yang paling banyak digunakan dalam bidang ekonomi dan bisnis dibandingkan dengan indeks lainnya.

8.4 PENYUSUNAN ANGKA INDEKS

Berdasarkan pengalaman, 4 persoalan pokok umumnya merupakan persoalan yang penting yang wajib diperhatikan dalam penyusunan indeks harga. Keempat persoalan tersebut acapkali menentukan mutu angka indeks.

a. Perumusan tentang tujuan penyusunan Indeks
   Sebelum data dikumpulkan dan pengukuran-pengukuran dilakukan, kita harus sudah dapat merumuskan apa yang diukur dan bagaimana cara pengukuran itu akan dilaksanakan. Tujuan penyusunan indeks adalah mengukur perubahan atau melakukan perbandingan antara peubahpeubah ekonomi maupun sosial.
b. Sumber dan syarat perbandingan data
   Mengenai syarat-syarat perbandingan bagi sebuah data, terdapat tiga hal yang harus diperhatikan:

1. dalam pembentukan indeks harga, tiap jenis barang harus memiliki kualitas yang kurang lebih sama selama periode perbandingan
2. Selalu dianjurkan untuk menggunakan data dari satu sumber. Data dari pelbagai sumber sukar sekali memenuhi syarat perbandingan, karena metode pengumpulan dan penyusunannya mungkin berbeda.

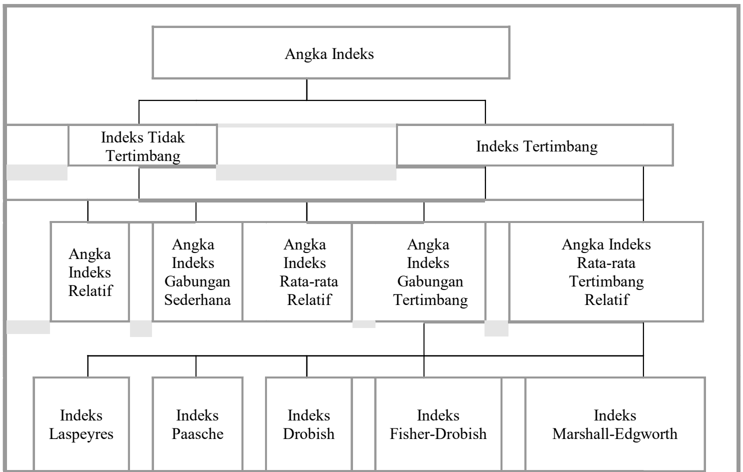
c. Pemilihan periode dasar Dalam pemilihan tahun dasar ada tiga ketentuan yang harus diperhatikan, yaitu:
   1. Sebagai tahun dasar, hendaknya dipilih tahun di mana keadaan perekonomian relatif stabil
   2. Tahun dasar sebagai dasar perbandingan hendaknya jangan terlalu jauh dari tahun-tahun yang hendak diperbandingkan
   3. Ada kalanya, kita ingin sekali mengukur kegiata atau perkembangan suatu peristiwa pada periode sesudah terjadi suatu kejadian atau perubahan yang penting.

d. Pemilihan timbangan (weight)

   Pada hakekatnya, timbangan mencerminkan betapa pentingnya suatu angka relatif terhadap angka-angka lain. Tanpa timbangan, angka-angka indeks kurang berguna bagi pengukuran perubahan maupun alat perbandingan.

8.5 METODE PERHITUNGAN ANGKA INDEKS

Menurut metode perhitungannya, angka indeks dibagi menjadi dua yaitu: (1) **Angka indeks tidak tertimbang**, dan (2) **Angka indeks tertimbang**. Angka indeks tidak tertimbang dibagi menjadi tiga yaitu; (1) Angka indeks agregatif sederhana, (2) Angka indeks relatif, dan (3) Angka indeks rata-rata hitung relatif. Sedangkan, angka indeks tertimbang dibagai dua yaitu; (1) Angka indeks agregatif, dan (2) Angka indeks rata-rata hitung relatif. merupakan indeks yang terdiri dari beberapa jenis barang (kelompok barang), misalnya indeks harga 9 (sembilan) macam bahan pokok, indeks ekspor Indonesia, indeks impor Indonesia, indeks biaya hidup, dan yang lainnya. Indeks agregatif memungkinkan untuk melihat persoalan secara keseluruhan, dan bukan melihat per individu. **Indeks Tertimbang** adalah indeks yang dalam penyusunannya telah mempertimbangkan faktor-faktor yang akan mempengaruhi naik turunnya angka indeks tersebut. Secara skematis angka indeks menurut metode perhitungannya dapat dinyatakan seperti Gambar berikut.



Gambar Skematis Angka Indeks Menurut Metode Perhitungannya

8.6 ANGKA INDEKS TIDAK TERTIMBANG

8.6.1 Angka Indeks Relatif

1. Angka Indeks Harga

Angka indeks harga adalah angka indeks yang memperhitungkan harga barang atau jasa dengan porsi yang sama tanpa memperhatikan bobot setiap barang dan jasa.

Rumus untuk indeks harga adalah sebagai berikut:

$$I_{t,0} = \frac{p_t}{p_0} x100\%$$

Dimana:

$I_{t,0}$ = Indeks harga pada waktu t

P_t = Harga pada waktu t

P_0 = Harga pada waktu dasar (0)

Contoh:

Tabel 8.1 Harga Barang menurut Jenis barang selama tahun 2004 – 2006

Jenis Barang	Harga		
	2004	**2005**	**2006**
A	100	150	200
B	200	250	300
C	500	600	700
D	400	500	600
Jumlah	1,200	1,500	1,800

Hitunglah indeks harga barang A pada tahun 2005 dan 2006 dengan tahun dasar 2004

Untuk tahun 2005:

$$I_{2005/2004} = \frac{P_{2005}}{P_{2004}} \text{ x } 100\% = \frac{150}{200} \text{ x } 100\% = 150\%$$

$$I_{2006/2005} = \frac{P_{2006}}{P_{2005}} \text{ x } 100\% = \frac{200}{100} \text{ x } 100\% = 200\%$$

2. Angka Indeks Kuantitas

Angka indeks kuatintas digunakan untuk melihat perkembangan kuantitas barang dan jasa. Indeks kuantitatif dihitung tanpa memberikan bobot setiap komoditi. Rumus angka indeks Kuantitas adalah sebagai berikut:

$$I_{t,0} = \frac{Q_t}{Q_0} \text{ x } 100\%$$

Dimana:

I_q = Indeks kuantitas relatif sederhana

q_t = Kuantitas tahun t

q_0 = Kuantitas tahun dasar (0)

Contoh:

TABEL 8.2 *Harga sayuran pada tahun 2004 – 2006*

Jenis Barang	Harga		
	2004	2005	2006
Kangkung	24,732	27,993	30,989
Bayam	1,551	1,659	1,785
Sawi	3,606	3,994	4,509
Brokoli	13,751	13,774	13,301

Dapatkan nilai indeks harga pada sayuran sawi tahun 2005 dan 2006. Gunakan tahun 2004 sebagai tahun dasar.

$$I_{2005/2004} = \frac{Q_{2005}}{Q} \text{ x } 100\% = \frac{3{,}994}{3{,}606} \text{ x } 100\% = 110.76\%$$

$$I_{2006/2005} = \frac{P_{2006}}{P_{2005}} \text{ x } 100\% = \frac{4{,}509}{3{,}606} \text{ x } 100\% = 125.04\%$$

8.6.2 Angka Indeks Gabungan Sederhana

Angka indeks untuk gabungan sederhana merupakan angka perbandingan untuk harga atau kuantitias nilai pada sekelompok barang dan jasa dengan mempunyai karakteristik kelompok yang sama. Contohnya adalah kelompok hasil perkebunan meliputi Teh, Kopi, dan Cengkeh, kelompok barang kebutuhan harian rumah tangga seperti beras, gula, garam, minyak sayur, ikan dan sayur. Untuk memudahkan indeksasi barang-barang tersebut dapat dikelompokkan ke dalam tiga bentuk indeks yaitu angka indeks harga, kuantitas dan nilai. Rumus untuk masing-masing indeks adalah sebagai berikut:

1. Angka Indeks Harga Gabungan Sederhana

Angka indeks ini digunakan untuk menghitung indeks barang dan jasa lebih dari satu. Di mana angka indeks ini menekankan pada agregasi barang dan jasa.

$$IA = \frac{\sum P_n}{\sum P_0} x100$$

Dimana :

P_n = harga tahun tertentu

P_0 = harga tahun dasar

$$IA = \frac{\sum P_n}{\sum P_0} x100$$

Dimana:

P_n = harga tahun tertentu

P_0 = harga tahun dasar

Contoh

Tabel 8.3 Harga eceran per satuan lima jenis barang per bulan di Kota Denpasar tahun 2009 dan tahun 2010

Jenis Barang	**Satuan**	**Harga Per Satuan (Rp)**	
		2009	**2010**
Minyak Goreng	Liter	12,518	12,529
Gula Pasir	Kg	8,355	10,628
Beras	Kg	5,505	6,541
Garam	Kg	3,241	3,333
Daging Ayam Ras	Kg	23,825	25,406

Sumber: BPS Provinsi Bali, 2011

Dari data di atas, dapatkan nilai indeks untuk harga agragatif tidak tertimbang bagi kelima barang. Berikan penjelasan pada hasil perhitungan yang didapat.

Jawaban
Tabel Akumulasi nilai harga barang kelima barang

Jenis Barang	Satuan	Harga (Rp)	
		2009	2010
Minyak Goreng	Liter	12,518	12,529
Gula Pasir	Kg	8,355	10,628
Beras	Kg	5,505	6,541
Garam	kg	3,241	3,333
DagingAyam Ras	Kg	23,825	25,406
Jumlah		53,444	58,437

Dari tabel dapat diketahui bahwa $\sum P_{2009}$= Rp 53.444 dan $\sum P_{2010}$ = 58,437
Selanjutnya per rumus (11.1) didapat:

$$P_{(n,0)} = \frac{\sum P_n}{\sum P_0}$$

$$P_{(n,0)} = \frac{58.437}{53.444}$$

$$P_{(n,0)} = 109.41$$

Jadi, indeks harga agregat (gabungan) tidak tertimbang kelima barang tersebut pada tahun 2010 dengan waktu dasar 2009 adalah 109.41. P(10,09) = 109.41 memiliki arti bahwa harga gabungan kelompok barang tersebut mengalami kenaikan sebesar (109.41-100) = 9.41% pada tahun 2010 dari harga gabungannya pada tahun 2009.

2. Angka Indeks Kuantitas Gabungan Sederhana

Angka indeks kuantitas agregat sederhana adalah angka indeks yang menunjukkan perbandingan antara jumlah kuantitas kelompok barang dan jasa pada periode tertentu dengan periode dasar. Rumus angka indeks kuantitas agregate sederhana adalah:

$$IKA = \frac{\sum Kt}{\sum K_0} x100$$

Dimana :

IKA = Indeks kuantitas agregat sederhana

Kt = kuantitas tahun t

K_0 = kuantitas tahun 0

Dimana:

IKA = Indeks kuantitas agregat sederhana

Kt = kuantitas tahun t

K_0 = kuantitas tahun 0

Contoh:

TABEL 8.4 ***Jumlah produksi sayuran pada Kabupaten X tahun 2010 dan 2011***

Jenis Barang	Banyaknya Produksi (Ton)	
	2010	2011
Bawang Merah	14,684	20,875
Bawang Putih	4,979	15,931
Bawang Daun	652	1,294
Kentang	2,261	5,107
Kubis	14,787	54,415
Kacang Panjang	120	100
Kacang Merah	159	17,051
Sawi	5,743	13,882
Terong	80	120

Sumber: Data Hipotetis

Hitunglah indeks rata-rata produksi (kuantitas) gabungan sayur mayur tersebut pada tahun 2011 degan waktu dasar tahun 2010

Jawaban

Jenis Barang	Banyaknya Produksi (Ton)	
	Q_0 (2010)	Q_n (2011)
Bawang Merah	14,684	20,875
Bawang Putih	4,979	15,931
Bawang Daun	652	1,294
Kentang	2,261	5,107
Kubis	14,787	54,415
Kacang Panjang	120	100
Kacang Merah	159	17,051
Sawi	5,743	13,882
Terong	80	120
Jumlah	43,465	128,775
Sumber: Data Hipotetis		

Berdasarkan tabel di atas, diperoleh nilai agregat pada $\sum Q_n = \sum Q_{11} = 128{,}775$, dan $\sum Q_0 = \sum Q_{10} = 43{,}465$.

$$Q_{(n,0)} = \frac{\sum Q_n}{\sum Q_0}$$

$$Q_{(11,10)} = \frac{128{,}775}{43{,}465}$$

$$Q_{(11,10)} = 296.27$$

$Q_{(11,10)}$ = 296,27, ini berarti produksi sayur mayur di kabupaten tersebut pada tahun 2011 mengalami kenaikan sebesar 196,27% dari tahun 2010.

3. Angka Indeks Nilai Gabungan Sederhana

Angka Indeks Nilai

Indeks nilai agregat (gabungan) beberapa barang menurut metode ini , dihitung dengan rumus sebagai berikut:

$$V_{(n,0)} = \frac{\sum V_n}{\sum V_0}$$

$V_{(n,0)}$ = Indeks nilai uang pada tahun n dengan tahun dasar 0
V_n = Nilai uang pada tahun n
V_o = Nilai uang pada tahun dasar 0

Contoh:

TABEL 8.5 ***Kuantitas dan Harga Barang Konsumsi di daerah X tahun 2010 dan 2011***

Jenis Barang	2010		2011	
	Harga/unit (Rp)	Kuantitas (Unit)	Harga/unit (Rp)	Kuantitas (Unit)
A	2,000	1	3,000	2
B	5,000	2	8,000	3
C	8,000	4	10,000	6
D	4,500	6	5,000	6

Sumber: data hipotetis

Dapatkan nilai indeks nilai gabungan keempat barang dia tas dengan menggunakan tahun 2010 sebagai tahun dasar.

Jawaban

Tabel Perhitungan Angka Indeks Nilai Gabungan EmpatJenis Barang

Jenis Barang	2010		2011		$V_0 = P_0 \times Q_0$	$V_n = P_n \times Q_n$
	P_0	Q_0	P_n	Q_n		
A	2,000	1	3,000	2	2,000	6,000
B	5,000	2	8,000	3	10,000	24,000
C	8,000	4	10,000	6	32,000	60,000
D	4,500	6	5,000	6	27,000	30,000
Jumlah					71,000	120,000

$$V_{(n,0)} = \frac{\sum V_n}{\sum V_0}$$

$$V_{(n,0)} = \frac{120{,}000}{71{,}000} \text{ x } 100$$

$$V_{(n,0)} = 169.01$$

$V_{(11,10)}$ = 169.01, memiliki arti bahwa nilai (dalam uang) ketiga jenis barang tersebut pada tahun 2011 naik sebesar (169.01-100) = 69.01% dari tahun 2010. Angka indeks ini merupakan hasil perhitungan indeks yang terdiri dari satu jenis barang saja. Misalnya indeks harga minyak goreng, indeks harga beras, indeks kuantitas beras, indeks kuantitas minyak goreng.

8.6.3 Angka Indeks Rata-Rata Relatif

1. Angka Indeks Harga

Menurut metode ini, indeks harga dapat dihitung dengan rumus;

$$P_{(n,0)} = \frac{1}{k} \sum \frac{P_n}{P_0} \text{ x } 100$$

Dimana,
k = banyaknya jenis barang

Contoh:

TABEL 8.6 ***Harga Eceran Rata-rata per Bulan 4 Bahan Pokok di Kota Denpasar Tahun 2009 dan 2010 (Dalam Rupiah).***

Jenis Bahan Pokok	Harga/satuan	
	2009	2010
Beras IR 64 (Kg)	5,505	6,541
Ikan Asin (Kg)	17,514	18,219
Minyak Goreng (Btl)	12,518	12,529
Gula Pasir (Kg)	8,355	10,625

Sumber: BPS Provinsi Bali, 2011

Hitunglah indeks rata-rata relatif harga eceran rata-rata empat bahan pokok tersebut pada tahun 2010 dengan waktu dasar tahun 2009.

Jawaban

Tabel Perhitungan Indeks Rata-rata Relatif Harga Eceran Empat Bahan Pokok

Jenis Bahan Pokok	**Harga/Unit**		**Pn/P0**
	(P_0)	**(P_n)**	
Beras IR 64 (Kg)	5,505	6,541	1,188
Ikan Asin (Kg)	17,514	18,219	1,040
Minyak Goreng (Btl)	12,518	12,529	1,001
Gula Pasir (Kg)	8,355	10,625	1,272
Jumlah			4,501

Dari Tabel diata, dapat diketahui bahwa $\sum \frac{P_n}{P_0} = 4.501$

$$P_{(n,0)} = \frac{1}{k}\sum \frac{P_n}{P_0} \text{ x } 100$$

$$P_{(10,9)} = \frac{1}{4}(4.501)\text{x } 100 = 112.525 \approx 112.53$$

Jadi, indeks rata-rata relatif harga eceran rata-rata per bulan 4 bahan pokok tersebut pada tahun 2010 dengan waktu dasar 2009 sebesar 112,53%. Ini berarti harga rata-rata per bulan keempat (4) bahan pokok tersebut mengalami kenaikan sebesar 12,53% dari harga eceran rata-ratanya pada tahun 2009.

Rumus untuk menghitung indeks kuantitas dan indeks nilai menurut metode rata-rata relatif, disesuaikan dengan rumus indeks harga, yaitu:

2. Angka Indeks Kuantitas

- Angka indeks kuantitas

$$Q_{(n,0)} = \frac{1}{k}\sum \frac{Q_n}{Q_0} \text{ x } 100$$

- Angka indeks nilai

$$V_{(n,0)} = \frac{1}{k}\sum\frac{V_n}{V_0} \text{ x } 100$$

8.7 ANGKA INDEKS TERTIMBANG

Indeks tertimbang merupakan angka indeks yang mencerminkan pentingnya suatu angka penimbang (bobot atau weight) terhadap angka–angka lainnya. Pemberian bobot angka penimbang tersebut ditentukan berdasarkan pentingnya barang/komoditi tersebut secara subyektif.

Sebagai faktor penimbang dalam perhitungan indeks tertimbang dipakai kuantitas barangnya. Pada umumnya timbangan yang dipakai adalah kuantitas barang yang dikonsumsi, diproduksi, dijual, dibeli, diekspor, diimpor dan atau yang lainnya. Di bawah ini hanya dibahas indeks harga saja.

Setiap barang dan jasa mempunyai tingkat utilitas yang berbeda, sehingga untuk menghitung angka indeks dimana banyak jenis komoditi perlu ada pembobotan pada setiap komoditi.

8.7.1 Angka Indeks Harga Gabungan

Angka indeks harga menurut metode ini dapat dihitung dengan rumus:

$$I_{(n,0)} = \frac{\sum P_n W}{\sum P_0 W} \text{ x } 100$$

W = Penimbang.
P_n = Harga barang pada tahun n.
P_o = Harga barang pada tahun dasar 0.

Indeks ini, merupakan angka indeks harga tertimbang, namuan pada dasarnya penyusunan indeks kuantitas tidak berbeda dengan penyusunan indeks harga. Bila penyusunan indeks harga berkisar pada perbandingan pn/p0 maka indeks kuantitas sebetulnya juga berkisar pada perbandingan qn/q0. Sebagai penimbangnya adalah harga pada tahun dasar (P_o). Sehingga angka indeks menurut metode ini dapat dihitung dengan rumus;

1. Indeks Laspeyres

$$I_L = \frac{\sum P_n Q_0}{\sum P_0 Q_0}$$

I_L = Indeks Laspeyres
Q_o = Kuantitas tahun dasar
P_n = Harga tahun tertentu
P_o = Harga tahun dasar

2. Indeks Paasche

Indeks ini, merupakan angka indeks tertimbang. Ssebagai penimbangnya adalah kuantitas pada tahun n (*Current Periode*). Sehingga angka indeks menurut metode ini dapat dihitung dengan rumus;

$$I_P = \frac{\sum P_n Q_n}{\sum P_0 Q_n}$$

I_P = angka indeks Paasche
P_n = harga tahun n
P_o = harga tahun dasar
Q_n = kuantitas tahun *n*

3. Indeks Irving Fisher

Menurut metode ini angka indeks dapat dihitung dengan rumus;

$$I_R = \sqrt{I_L \text{ x } I_P}$$

$$IF = \sqrt{\left[\frac{\sum P_n q_0}{\sum P_0 q_0} \; X\, 100\right] X \left[\frac{\sum P_n q_n}{\sum P_0 q_n} \; X\, 100\right]}$$

I_F = angka indeks Irving Fisher
I_L = angka indeks Laspeyres
I_P = angka indeks Paasche

4. Angka Indeks Drobisch

Metode indeks Drobisch dapat dihitung dengan menggunakan rumus seperti berikut:

$$I_D = \frac{I_L + I_P}{2}$$

atau

$$ID = \frac{\left[\dfrac{\sum P_n q_0}{\sum P_0 q_0} X\,100\right] + \left[\dfrac{\sum P_n q_n}{\sum P_0 q_n} X\,100\right]}{2}$$

5. Indeks Marshall-Edgeworth

Metode ini memakai ($Q_o + Q_n$) sebagai faktor penimbang, dan menurut matode ini angka indeks dihitung dengan rumus;

$$I_M = \frac{\sum P_n(\,Q_0 + Q_n)}{\sum P_0(Q_0 + Q_n)} \text{ x } 100$$

I_M = angka indeks Marshall-Edgeworth

Contoh:
Data mengenai harga dan kuantitas produksi empat jenis barang di Provinsi "X" disajikan dalam Tabel 8.7 berikut:

TABEL 8.7 ***Harga dan Kuantitas Produksi Empat Jenis Barang di Provinsi " X "Tahun 2010–2011***

Jenis Barang	Harga/Unit (Rp)		Kuantitas Produksi (Unit)	
	2010	2011	2010	2011
A	500	525	2	4
B	800	900	5	6
C	600	700	3	4
D	300	400	10	15

Hitunglah indeks harga agregatif tertimbang barang-barang tersebut pada tahun 2011 dengan tahun dasar 2010 (a) Dengan metode Laspeyres.

(b) Dengan metode Paasche.

(c) Dengan metode Irving Fisher.

(d) Dengan metode Drobish.

(e) Dengan metode Marshall-Edgeworth.

Jawaban

(a) Indeks harga Laspeyres

Tabel Perhitungan Angka Indeks Harga Laspeyres Barang Tahun 2011 dengan Tahun Dasar 2010

Jenis Barang	Harga/Unit (Rp)		Kuantitas (Unit)			
	2010	2011	2010	2011		
	(P_o)	(P_n)	(Q_o)	(Q_n)	$P_n.Q_o$	$P_o.Q_o$
A	500	525	2	4	1,050	1,000
B	800	900	5	6	4,500	4,000
C	600	700	3	4	2,100	1,800
D	300	400	10	15	4,000	3,000
Jumlah					11,650	9,800

Dari Tabel di atas, diketahui $\sum P_n Q_0 = 11.650$ dan $\sum P_0 Q_0 = 9.800$

$$I_L = \frac{\sum P_n Q_0}{\sum P_0 Q_0}$$

$$I_L = \frac{11,650}{9,800}$$

$$I_L = 118.88$$

(b) Indeks harga Paasche

Tabel Perhitungan Angka Indeks Harga Paasche Barang Tahun 2011 dengan Tahun Dasar 2010

Jenis Barang	Harga/Unit (Rp)		Kuantitas (Unit)			
	2010	2011	2010	2011		
	(P_o)	(P_n)	(Q_o)	(Q_n)	Pn . Qn	Po . Qn
A	500	525	2	4	2,100	2,000
B	800	900	5	6	5,400	4,800
C	600	700	3	4	2,800	2,400
D	300	400	10	15	6,000	4,500
Jumlah					16,300	13,700

Dari Tabel diketahui $\sum P_n Q_n$ = 16,300 dan $\sum P_0 Q_n$= 13,700

Maka, perhitungan dengan rumus dapat dilakukan seperti berikut:

$$I_P = \frac{\sum P_n Q_n}{\sum P_0 Q_n}$$

$$I_P = \frac{16{,}300}{13{,}700}$$

$$I_P = 118.97$$

(c) Indeks harga Irving Fisher

$$I_R = \sqrt{I_L \text{ x } I_P}$$

$$I_R = \sqrt{187.90 \text{ x } 118.97}$$

$$I_R = 118.92$$

(d) Indeks harga Drobisch

$$I_D = \frac{I_L + I_P}{2}$$

$$I_D = \frac{118.88 + 118.97}{2}$$

$$I_D = 118.92$$

(e) Indeks harga Marshall - Edgeworth
Tabel Perhitungan Angka Indeks Harga Marshall-Edgeworth Tahun 2011 dan tahun dasar 2010

Jenis Barang	**Harga/Unit (Rp)**		**Kuantitas (Unit)**				
	2010	**2011**	**2010**	**2011**		**P_n x**	**P_o x**
	(P_o)	**(P_n)**	**(Q_o)**	**(Q_n)**	**Qo+Qn**	**(Q_o+Q_n)**	**(Q_o+Q_n)**
A	500	525	2	4	6	3,150	3,000
B	800	900	5	6	11	9,900	8,800
C	600	700	3	4	7	4,900	4,200
D	300	400	10	15	25	10,000	7,500
Jumlah						27,950	23,500

Dari hasil perhitungan di atas, diketahui nilai nilai $\sum P_n (Q_0 + Q_n) = 27{,}950$ dan $P_0 (Q_0 + Q_n) = 23{,}500$

$$I_M = \frac{\sum P_n(Q_0 + Q_n)}{\sum P_0(Q_0 + Q_n)} \text{ x } 100$$

$$I_M = \frac{27{,}950}{23{,}500} \text{ x } 100$$

$$I_M = 118.93$$

Hasil perhitungan menurut Laspeyres biasanya lebih besar dibandingkan hasil perhitungan Paasche. Hal ini terjadi, bila penimbanganya adalah konsumsi masyarakat. Angka indeks Laspeyres hasil perhitungannya cenderung *over estimate* (berlebihan ke atas) dan angka indeks Paasche hasil perhitungannya cenderung *under estimate* (berlebihan ke bawah). Untuk mengatasi hal-hal tersebut, Irving Fisher, Drobish dan Bowley mengambil jalan tengah, yaitu dengan jalan mengambil nilai rata-rata dari indeks Laspeyres dan indeks Paasche. Irving Fisher mengambil rata-rata ukur dari indeks Laspeyres dan Paasche, sedangkan Drobish dan Bowley mengambil rata-rata hitung dari indeks Laspeyres dan Paasche.

Indeks menurut Fisher ini secara teoritis merupakan indeks yang paling baik, maka dari itu indeks Fisher sering disebut sebagai **Fisher Ideal indeks Numbers.** Walaupun demikian, dalam praktek indeks menurut Laspeyres lah yang sering digunakan, mengingat untuk menghitungnya hanya cukup dengan mencari P_n saja, sedangkan P_o dan Q_o angkanya konstan. Tidak demikian halnya dengan indeks lainnya seperti Paasche dan Fisher.

8.7.2 Angka Indeks Harga Rata-Rata Tertimbang Relatif

Indeks harga sekelompok barang menurut metode ini akan dihitung de ngan rumus;

$$P_n = \frac{\sum(P_n / P_0)\, W}{\sum W} \text{ x } 100$$

W = menunjukkan nilai timbangan
P_n = harga barang pada tahun n
P_o = harga barang pada tahun dasar 0

(a) Jika dipakai faktor penimbang nilai pada tahun dasar, maka rumus tersebut menjadi;

$$P_n = \frac{\sum(P_n / P_0)\,(P_0 Q_0)}{\sum P_0 Q_0} \text{ x } 100$$

(b) Jika dipakai faktor penimbang nilai pada tahun n maka rumus tersebut menjadi:

$$P_n = \frac{\sum(P_n / P_0)\,(P_n Q_n)}{\sum P_n Q_n} \text{ x } 100$$

Contoh:

TABEL 8.8 ***Perhitungan Angka Indeks Harga Tertimbang Rata-rata Relatif Atas Empat Jenis Barang***

Jenis Barang	Harga/Unit (Rupiah)		Kuantitas (Unit)		Pn/Po	Po.Qo	Pn.Qn	Pn/Po x Po.Qo	P_n/P_o x Pn.Qn
	2010 P_o	2011 P_n	2010 Q_o	2011 Q_n					
A	500	525	2	4	1,050	1.000	2.100	1.050	2.205
B	800	900	5	6	1,125	4.000	5.400	4.500	6.075
C	600	700	3	4	1,167	1.800	2.800	2.101	3.268
D	300	400	10	15	1,333	3.000	6.000	3.999	7.998
Jumlah						9.800	16.300	11.650	19.546

Sumber*:* Data hipotetis

1) Bila dipakai timbangan nilai pada tahun dasar 2010 maka angka indeksnya:

$$P_n = \frac{\sum (P_n / P_0)\ (P_0 Q_0)}{\sum P_0 Q_0} \text{ x } 100$$

$$P_n = \frac{11.650}{9.800} \text{ x } 100$$

$$P_n = 118.87$$

2) Bila dipakai timbangan nilai pada tahun n (2011) maka angka indeksnya:

$$P_n = \frac{\sum (P_n / P_0)\ (P_n Q_n)}{\sum P_n Q_n} \text{ x } 100$$

$$P_n = \frac{19.546}{16.300} \text{ x } 100$$

$$P_n = 119.91$$

8.8 ANGKA INDEKS RANTAI

Angka indeks berantai adalah angka indeks yang menggunakan waktu dasar selalu satu tahun sebelum tahun yang dihitung angka indeksnya. Misalnya angka indeks tahun 2009 dihitung dengan memakai tahun dasar 2008, angka indeks tahun 2010 dihitung dengan memakai tahun dasar 2009, demikian seterusnya. Berikut ini akan dipelajari dua macam indeks berantai yaitu: (1) Indeks harga relatif berantai, dan (2) Angka indeks tertimbang berantai. Perubahan angka indeks yang mempunyai waktu dasar tertentu. Waktu dasar tersebut tetap selama pembuatan indeks dari tahun ke tahun. Tetapi sering kali seseorang membutuhkan waktu dasar yang berubahubah. Perubahan dapat dilakukan setiap satu waktu, dua waktu atau lebih dari dua waktu. Bila satu satuan waktu (tahun, bulan) maka simbolnya $I_{t,t-}$ 1, 2 satuan waktu $I_{t,t-2}$ dst.

8.8.1 Angka Indeks Harga Relatif Berantai

Perubahan angka indeks yang mempunyai waktu dasar tertentu. Waktu dasar tersebut tetap selama pembuatan indeks dari tahun ke tahun. Tetapi sering kali seseorang membutuhkan waktu dasar yang berubah-ubah. Perubahan dapat dilakukan setiap satu waktu, dua waktu atau lebih dari dua waktu. Bila satu satuan waktu (tahun, bulan) maka simbolnya It,t – 1, 2 satuan waktu It,t-2 dst. Rumus yang digunakan untuk mencari indeks berantai adalah:

$$I_{t,t-1} = \frac{P_t}{P_{t-1}} \text{ x } 100$$

Dimana:

P_t = Harga tahun t

P_{t-1} = Harga tahun t – 1

Bila yang akan dihitung angka indeks berantai dari kuantitas. Rumus yang digunakan adalah:

$$I_{t,t-1} = \frac{Q_t}{Q_{t-1}} \text{ x } 100$$

Di mana:

Q_t = kuantitas tahun t

Q_{t-1} = kuantitas tahun t – 1

Contoh:

TABEL 8.9 ***Nilai Eksporsr Barang X tahun 1990 - 1993***

Tahun	**Ekspor Barang X (Ribuan)**
1990	253
1991	345
1992	487
1993	679

Buatlah indeks berantai untuk tahun 1991, 1992 dan 1993 dengan waktu dasar satu tahun sebelumnya.

Jawaban

$$I_{1991,1990} = \frac{q_{1991}}{q_{1990}} X\,100 = \frac{345}{253} X\,100 = 136.36$$

$$I_{1992,1991} = \frac{q_{1992}}{q_{1991}} X\,100 = \frac{487}{345} X\,100 = 141.12$$

$$I_{1993,1992} = \frac{q_{1993}}{q_{1992}} X\,100 = \frac{679}{487} X\,100 = 139.42$$

Keuntungan menggunakan angka indeks berantai adalah:

a. Kita dapat memasukkan komoditi baru atau menambah komposisi komoditi.

b. Indeks berantai yang telah kita peroleh dapat dipergunakan untuk membuat indeks pada tahun tertentu dengan dasar waktu tetap. Rumus yang digunakan adalah:

$$I_{t+1,t-1} = (I_{t,t-1})\,(I_{t+1,t})$$

Dengan memakai contoh di atas, maka dapat diketahui indeks pada tahun 1992 dan 1993 dengan tahun dasar 1990 adalah:

$$I_{t+1,t-1} = (I_{t,t-1})\ (I_{t+1,t})$$

$$I_{1992,1990} = (1.363)(1{,}411) \text{ x } 100 = 192.5$$

$$I_{1993,1990} = (1.925)(1{,}394) \text{ x } 100 = 268.4$$

Contoh

TABEL 8.10 ***Angka Indeks Harga Relatif Biasa dan Angka Indeks Relatif Berantai Harga Rata-rata Beras per Bulan Tahun 2007- 2010***

Tahun	Harga (Rp/Kg)	Indeks Biasa	Indeks Berantai
2007	4,900	100 (Th. dasar)	100
2008	5,154	105.18	105.18
2009	5,505	112.34	106.81
2010	6,541	133.49	118.82

Angka indeks biasa pada kolom 3 (tiga) didapat sebagai berikut;

$$P_{(08,07)} = \frac{5.154}{4.900} \text{x } 100 = 105.18$$

$$P_{(09,07)} = \frac{5.505}{4.900} \text{x } 100 = 112.34$$

$$P_{(09,07)} = \frac{6.541}{4.900} \text{x } 100 = 133.49$$

Angka indeks berantai pada kolom 4 (empat) didapat sebagai berikut;

$$P_{(08,07)} = \frac{5.154}{4.900} \text{x } 100 = 105.18$$

$$P_{(09,08)} = \frac{5.505}{5.154} \text{x } 100 = 106.81$$

$$P_{(10,09)} = \frac{6.541}{5.505} \text{x } 100 = 118.82$$

Perhatikan Tabel 11.12, dapat diketahui bahwa:
Pada indeks biasa, angka indeks yang dicari ***selalu*** dibandingkan dengan tahun dasar (tahun 2007).

Tahun dasarnya tetap.
Tahun 2008 harga per kg beras naik 5,83% dari harga pada tahun 2007.
Tahun 2009 harga per kg beras naik 12,35% dari harga pada tahun 2007.
Tahun 2010 harga per kg beras naik 33,49% dari harga pada tahun 2007.

Pada indeks berantai, angka indeks yang dicari selalu dibandingkan dengan satu periode waktu dari waktu yang akan dihitung angka indeksnya, sehingga kenaikan harga setiap tahun dapat diketahui.

Tahun 2008 harga per kg beras naik 5,83% dari harga pada tahun 2007.
Tahun 2009 harga per kg beras naik 6,81% dari harga pada tahun 2008.
Tahun 2010 harga per kg beras naik 18,82% dari harga pada tahun 2009.

Hubungan indeks relatif secara berantai untuk beberapa tahun dapat dinyatakan sebagai berikut:

(1) Tiga indeks berantai

$$P_{(n,\ n-3)} = \frac{P_{(n-2)}}{P_{(n-3)}} \text{x} \frac{P_{(n-1)}}{P_{(n-2)}} \text{x} \frac{P_{(n)}}{P_{(n-1)}}$$

(2) Empat indeks berantai

$$P_{(n,\ n-4)} = \frac{P_{(n-3)}}{P_{(n-4)}} \text{x} \frac{P_{(n-2)}}{P_{(n-3)}} \text{x} \frac{P_{(n-1)}}{P_{(n-2)}} \text{x} \frac{P_{(n)}}{P_{(n-1)}}$$

(3) Sedangkan untuk data pada Tabel 10.15 dapat dinyatakan sebagai berikut:

$P_{(87,84)} = P_{(08,07)}$ x $P_{(09,\ 08)}$ x $P_{(10,\ 09)}$

$$P_{(87,84)} = \frac{P_{(08)}}{P_{(07)}} \text{x} \frac{P_{(09)}}{P_{(08)}} \text{x} \frac{P_{(10)}}{P_{(09)}}$$

8.8.2 Angka Indeks Tertimbang Berantai

Dengan metode ini angka indeks dapat dihitung dengan rumus:

$$P_{(n,n-1)} = \frac{\sum P_n W}{\sum P_{(n-1)} W} \text{ x } 100$$

Contoh:

TABEL 8.11 ***Harga Eceran Tiga Jenis Bahan Kebutuhan Pokok Tahun 2008 - 2010***

Jenis Barang	Harga/kg			Timbangan W
	2008	2009	2010	
Beras	5,154	5,505	6,541	10
Gula Pasir	6,284	8,355	10,628	6
Garam	2,816	3,241	3,333	4

Sumber: BPS Provinsi Bali, 2011

Jawaban

Maka, $P_{(09,08)}$ untuk dapat dihitung seperti berikut:

$$P_{(n,n-1)} = \frac{\sum P_n W}{\sum P_{(n-1)} W} \text{ x } 100$$

$$P_{(09,08)} = \frac{\sum P_{09} W}{\sum P_{08} W} \text{ x } 100$$

$$P_{(09,08)} = \frac{(5{,}505 \,.\, 10) + (8{,}355 \,.\; 6) + (3{,}241 \,.\; 4)}{(5{,}154 \,.\, 10) + (6{,}284 \,.\; 6) + (2{,}816 \,.\; 4)} \text{ x } 100$$

$$P_{(09,08)} = 117.54$$

Sedangkan untuk $P_{(10,09)}$ seperti berikut:

$$P_{(n,n-1)} = \frac{\sum P_n W}{\sum P_{(n-1)} W} \text{ x } 100$$

$$P_{(10,09)} = \frac{\sum P_{10}W}{\sum P_{09}W} \text{ x } 100$$

$$P_{(09,08)} = \frac{(6{,}541\,.\,10) + (10{,}628\,.\;6) + (3{,}333\,.\;4)}{(5{,}505\,.\,10) + (8{,}355\,.\;6) + (3{,}241\,.\;4)} \text{ x } 100$$

$$P_{(09,08)} = 120.31$$

Penentuan timbangan (*weigh*t) dalam perhitungan bersifat subyektif, jika penentuan timbangan dilakukan secara obyektif, maka besar timbangan (*weight*) disesuaikan dengan kuantitas barangnya.

8.9 ANGKA INDEKS UNTUK PROSES DEFLASI

Upah nominal yang tinggi tidak selalu mencerminkan tingkat hidup yang lebih baik dari keadaan sebelumnya apabila perkembangan tingkat harga barang-barang kebutuhan pokok sehari-hari (biaya hidup) tinggi pula. Seorang buruh atau pegawai (karyawan tetap) lebih senang menerima gaji yang lebih kecil dengan daya beli besar dari pada gaji yang lebih besar tetapi dengan daya beli kecil. Dengan kata lain, seorang buruh atau pegawai akan lebih senang menerima upah nyata (daya beli) dari uang tersebut dari pada upah uang (nilai nominal) dari uang yang diterima. Besar kecilnya upah nyata ini, tergantung dari indeks biaya hidup (*cost of living index*) atau indeks harga konsumen (*consumer's price index*).

Indeks harga konsumen tidaklah sama dengan indeks biaya hidup. Indeks harga konsumen disusun berdasarkan harga-harga sekelompok barang atau jasa tanpa memasukkan semua jenis biaya, seperti bermacam-macam pajak. Selain itu sebagian biaya hidup lebih ditentukan oleh selera atau gaya hidup dibanding harga.

Sampai saat ini Badan Pusat Statistik (BPS) belum menerbitkan indeks biaya hidup yang diterbitkan adalah indeks harga konsumen (IHK) tahunan dan bulanan. Bila indeks biaya hidup tidak tersedia, maka indeks harga konsumen sering digunakan sebagai pengganti indeks biaya hidup. IHK dihitung dengan rumus Laspeyres.

Dengan demikian untuk menghitung upah nyata (upah riil) dengan proses deflasi, sebagai deflator dapat digunakan IBH atau IHK dengan rumus masing-masing berikut ini:

$$\text{Upah riil} = (\text{Upah Nominal}) / (\text{Indeks Biaya Hidup X } 100)$$

Contoh:

TABEL 8.12 ***Perhitungan Upah Riil Tahun 2007-2011***

Tahun	Upah Nominal/hari (Rupiah)	Indeks Harga Konsumen (2007 =100)	Upah Riil (Rupiah)
2007	400.000	100	400.000
2008	500.000	80	625.000
2009	600.000	125	480.000
2010	750.000	200	375.000
2011	800.000	320	250.000

Sumber: Data hipotetis

Dengan mengaplikasikan rumus, perhitungan upah nyata (riil) dari tahun 2007 sampai dengan 2011 masing-masing dapat dilakukan sebagai berikut:

$$\text{Upah nyata tahun 2008} = \frac{500.000}{80} \text{ x } 100 = 625.000$$

$$\text{Upah nyata tahun 2009} = \frac{600.000}{125} \text{ x } 100 = 480.000$$

$$\text{Upah nyata tahun 2010} = \frac{750.000}{200} \text{ x } 100 = 375.000$$

$$\text{Upah nyata tahun 2011} = \frac{800.000}{320} \text{ x } 100 = 250.000$$

Upah nyata/riil dari tahun 2007 hingga tahun 2011 dimuat pada kolom terakhir Tabel 11.16. Dari Tabel 11.16 dapat diketahui bahwa upah riil dari tahun 2007 hingga 2011 nilainya berfluktuasi. Pada tahun 2008 mengalami kenaikan, sedangkan pada tahun 2009 hingga 2011 mengalami penurunan.

Bab 9
Peluang

9.1 PENGANTAR

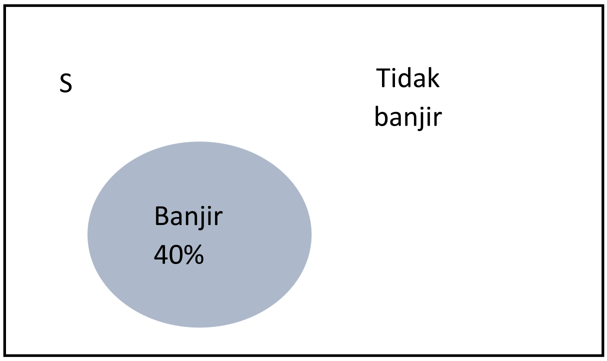
Setiap keputusan yang diambil oleh seseorang pasti akan menimbulkan suatu resiko, hal ini bergantung dengan seberapa jauh keputusan diambil. Dalam pengambilan keputusan, tidak terlepas dari nilai-nilai peluang yang akan didapatkan, baik itu akan mengalami resiko, ataupun kepastian yang akan menghasilkan hasil yang baik. Sebagai contoh, ada suatu industri makanan, yang mengambil keputusan yang tidak mengikuti tren pasar, hanya meyakini dan mempercayai nilai-nilai peluang dalam mengambilan keputusan, maka akan berakibat fatal bagi industry makanan tersebut. Salah satu alat analisis statistic yang digunakan untuk memecahkan suatu masalah adalah konsep probabilitas. Probabilitas sering disebut sebagai peluang, yaitu suatu peristiwa yang kemungkinan bisa terjadi atau tidak.

9.2 PENGERTIAN PELUANG ATAU PROBABILITAS

Pengertian Probabilitas secara umum disebut peluang, suatu kebolehjadian pada suatu peristiwa yang kemungkinan bisa terjadi, yang dinyatakan sebagai P(E). Definisi probabilitas bisa dikatakan terkadinya peristiwa n_i yang dihasilkan dari suatu semesta kosmos atau ruang contoh S dalam percobaan tertentu yang dinyatakan sebagai P(E). Secara kuantitatif, nilai peluang/ probabilitas P(E) merupakan bentuk bilangan pecahan antara 0 dan 1. Saat peluang 0 maka kemungkinan adanya suatu kejadian/peristiwa, tidak mungkin terjadi. sedangkan saat peluang 1 kemungkinan peristiwa bisa terjadi. Menurut Soedibjo (2010:1) adalah suatu cara untuk menyatakan kesempatan terjadinya peristiwa. Secara kualitatif, peluang dinyatakan dalam bentuk kata sifat untuk menunjukkan kemungkinan terjadinya suatu keadaan, seperti lemah, kuat atau baik.

Bisa dicontohkan, mata uang logam yang dilemparkan, pada hasil lemparan tersebut kemungkinan kan keluar gambar kepala dan ekor. Secara teori probabilitas, dua sisi tersebut mempunyai kemungkinan yang sama yaitu $^1/_2$ dalam 1 kali memunculkan gambar kepala maupun ekor pada saat dua kali lemparan. Atau bisa terjadi saat seseorang mengunjungi tempat pembelanjaan, dan membeli sekitar 5 baju dengan harga Rp. 100.000 per baju, maka bisa dipastikan dengan mudah, harga yang harus dibayar seharga Rp. 500.000. namun bisa terjadi saat seseorang ingin membeli baju namun membawa uang yang tidak pasti, karena belum memastikan berapa harga baju tersebut, hanya mengandalkan suatu peluang harga yang dipikirkannya, maka saat membeli baju bisa terjadi kurangnya uang untuk membeli baju.

Sebagai contoh lain saat mendengar akan terjadi peristiwa banjir dalam bentuk peluang, baik dalam bentuk kualitatif bahwa terjadi kemungkinan kecil bahwa akan terjadi banjir pada esok hari. Dan dalam bentuk kuantitatif, bahwa kemungkinan akan terjadi banjir sekitar 40%. Saat semakin banyaknya nilai peluang pada suatu perhitungan, maka semakin besar seseorang menyakini persitiwa tersebut akan terjadi. Gejala sebuah peristiwa tidak hanya dikaji dari satu sisi, namun bisa juga karena pengaruh waktu, jika melibatkan variable pada peristiwa tersebut. Peluang didasarkan oleh latar belakang ilmiah. Salah satu cara untuk menyatakan peluang suatu kejadian adalah Diagram Venn.



GAMBAR 9.1 ***Diagram Venn***

Diagram Venn berbentuk persegi Panjang untuk menyatakan semua kejadian bisa terjadi dan lingkaran berfungsi untuk menggambarkan peluang terjadinya kejadian atau peristiwa tersebut. Diagram venn pada umumnya tidak menggunakan skala yang sesungguhnya, artinya jika peluang terjadi peristiwa hujan 30% bukan berarti lingkaran yang dimaksudkan luasnya 30% dari luas persegi Panjang. Oleh karena itu, Peluang atau probabilitas adalah suatu kejadian yang dapat mengungkapkan suatu kepercayaan atau pengetahuan tentang peluang terjadinya kejadian atau peristiwa tersebut akan terjadi atau tidak terjadi. Suatu eksperimen dalam peluang digunakan untuk mendapatkan suatu penelitian dalam percobaan dan hasilnya dapat dipastikan atau diramalkan. Percobaan yang sederhana ini adalah menghitung peluang dadu, koin atau mata uang. Selain itu, membahas peluang tidak bisa lepas dari percobaan, ruang sampel dan kejadian atau peristiwa.

Menurut Definisi Klasik, peluang didefinisikan sebagai perbandingan antara banyaknya peristiwa E yang terjadi (n) dengan banyaknya kemungkinan terjadinya seluruh peristiwa (N) yang saling eksklusif.

$$P(E): \frac{n}{N} \tag{9.1}$$

Misalnya kotaak berisi 50 kelereng yang identik kecuali warnanya. Kelereng tersebut terdiri dari 20 kelereng berwarna merah, 6 kelereng berwarna kuning, dan 10 kelereng berwarna biru. Selanjutnya kelereng tersebut diaduk dandiambil satu kelereng dari dalam kotak. Peluang kelereng terambil masing-masing:

$P(\text{merah}) \frac{20}{50} = 0{,}4$

$P(\text{kuning}) \frac{6}{50} = 0{,}12$

$P(\text{biru}) \frac{10}{50} = 0{,}2$

Menurut Definisi Empirik , peluang didefinisikan sebagai limit dari frekuensi relative n terhadap N apabila jumlah pengamatan diperbanyak dampai tak terhingga. Misalkan produksi padi diperiksa 2000 dan terdapat barang yang sudah rusak sebanyak 66. Frekuensi relative pada kerusakan produksi sebesar 0,033. Kemudian, diperiksa lagi, terdapat 4000 produksi padi dan terjadi kerusakan sebesar 150. Frekuensi relatifnya sebesar 0,0375.

9.3 PENDEKATAN PROBABILITAS

Terdapat tiga tingkat pendekatan probabilitas, yaitu:

9.3.1 Pendekatan Klasik

Pendekatan klasik didasarkan pada peristiwa yang bersifat eksklusif dan asumsi hasil dari suatu percobaan, masing-masing akan mempunyai kesempatan yang sama untuk muncul (*equally*). Menurut pendekatan klasik, saat terjadi peristiwa (E) bisa dinyatakan sebagai rasio satu kejadian dari seluruh kejadian, saat suatu kejadian atau peristiwa mempunyai kesempatan yang sama.

$$P(E) : \frac{n\,(E)}{n\,(S)} \qquad (9.2)$$

Keterangan:
P(E) : Peluang A
n(E) : Banyaknya kejadian A
n(S) : Banyaknya anggota ruang sampel A

Contoh soal 1:
Sebuah dadu dilemparkan sekali. Dadu tersebut berisi 6 sisi. Berapakah peluangnya, jika terdapat 2 sisi dadu yang akan muncul?

Banyaknya suatu titik sampel; n(s) = 6, jika muncul 2 sisi maka ⅙ = 0,167

Contoh soal 2:

Terdapat 200 mahasiswa di suatu Universitas di Jakarta, terdapat 50 orang diantaranya Wanita. Berapakah peluang mahasiswa Wanita tersebut?

$$\text{P(E)} : \frac{50}{200} = \frac{1}{4}$$

9.3.2 Pendekatan Relatif

Pendekatan ini didasarkan oleh frekuensi relatif. Berikut rumus pendekatan relative:

$$\text{P(E)} : \frac{ni}{s} \tag{9.3}$$

Contoh soal 1:

Sebuah Penelitian yang dilakukaan oleh Lembaga survei dimana terdapat 800 ribu orang lulus dari Universiyas fakultas hukum. Penelitian menunjukkan sebuah percobaan, hasilnya terdapat 450 dari 800 ribu orang tidak bekerja sesuai bidangnya. Berapakah probabilitas seorang lulusan fakultas hukum tidak bekerja sesuai bidangnya?

Probabilitas yang akan terjadi: $\frac{450}{800}$ = 0,56

Kesimpulannya, berdasarkan pengalaman masa lalu probabilitas bagi mahasiswa yang tidak bekerja sesuai bidang studinya adalah 0,56.

Contoh soal 2:

Sebuah Perusahaan elektronik memproduksi Hp sebanyak 500.000 unit merk tertentu, namun terdapat 50 unit barang yang cacat. Maka bisa disimpulkan, peluang Hp yang cacat adalah 50/500.000 = 0,0001. Maka rata-rata persentase barang yang cacat diperkirakan 0,0001.

9.3.3 Pendekatan Subyektif

Jika terjadi suatu kejadian atau peristiwa hanya terjadi beberapa kali atau tidak ada informasi yang relative, maka probabilitas hanya ditentukan menggunakan pengetahuan seseorang, atau keyakinan individu. Dengan adanya probabilitas yang tidak disertai dengan bukti empiris atau fakta maka disebut dengan probabilitas subyektif. Karena bersifat individu, nilai probabilitas dari satu individu ke individu akan berbeda walaupun peristiwa atau kejadian bersifat sama. Pendekatan ini biasanya digunakan oleh para peneliti yang berpengalaman untuk meramalkan suatu kejadian.

9.4 KEJADIAN MAJEMUK

9.4.1 Kejadian Saling Lepas

Kejadian yang bersifat saling lepas (*mutually exclusive*) merupakan probabilitas yang terjadi jika kejadian E atau kejadian F saling lepas satu dengan yang lain atau tidak mungkin terjadi keduanya saat terjadi bersama-sama. Bisa diartikan, jika suatu peristiwa E terjadi, maka F tidak akan terjadi atau tidak ada peristiwa lain disaat yang sama. Jika Probabilitas kejadian dilambangkan dengan $\mathrm{P}\ (\mathrm{E} \cap \mathrm{F}) = 0$.

Berdasarkan hukum penjumlahan:

$$\mathrm{P}\ (\mathrm{E} \cup \mathrm{F}) = (\mathrm{E}) + \mathrm{P}\ (\mathrm{F}) \qquad (9.4)$$

Jika E_1, E_2, …. E_n saling lepas, maka probabilitas kejadian tersebut adalah:

$$P(E_1 \cup E_2\ \cup, \ldots E_\mathrm{n}) = P(E_1) + P(E_2) + \cdots + P(E_\mathrm{n}) \qquad (9.5)$$

Bila kejadian-kejadian E dan E′ merupakan komplemen satu dengan lainnya, maka $\mathrm{E}\ \cap\ \mathrm{E}' = \phi$ dan $\mathrm{E}\ \cup\ \mathrm{E}' = \mathrm{S}$

Contoh:
Terdapat kartu yang membentuk kata:

A U S T R A L I A

Dari 9 kartu diambil 1 kartu secara acak. Berapakah probabilitas terambilnya kartu A atau kartu S.

Probabilitas terambil kartu A, $\mathrm{P(A)} = \frac{3}{9}$

Probabilitas terambil kartu S, $\mathrm{P(S)} = \frac{1}{9}$ akan tetapi $\mathrm{P(A \cap S)} = \phi$.

Sehingga, $\mathrm{P(A \cup S)} = \mathrm{P(A)} + \mathrm{P(S)} = \frac{3}{9} + \frac{1}{9} = 0{,}444$.

Contoh:
Terdapat Uang logam dilemparkan sebanyak 8 kali selama berturut-turut. Berapakah probabilitas sekurang-kurangnya gambar garuda muncul selama 1 kali?

S mempunyai $2^8 = 256$. Dimisalkan A merupakan kejadian yang ditanyakan sehingga, P(A) = 1 – P(A′) dengan A′ menyatakan kejadian bahwa sisi gambar garuda tidak pernah muncul dengan

$$P(A') = \frac{1}{256}$$

$$P(A) = 1 - P(A')$$

$$= 1 - \frac{1}{256} \quad = 0,996$$

Contoh:
Apabila sebuah dadu bermata 6 dilempar, maka berapa peluang untuk tidak mendapat sisi dadu 4?

Jawaban:
Ada enam mata dadu, dengan sisi dadu 4 berjumlah satu maka n(S) = 6 dan n(K) = 1
$P(dadu) = \frac{n(K)}{n(S)} = \frac{1}{6}$, sehingga peluang komplemen dari kejadian tersebut adalah
$P(dadu') = 1 - P\ (dadu)$

$P(dadu') = 1 - \frac{1}{6}$

$P\ (dadu') = \frac{5}{6}$, Jadi peluang untuk tidak mendapatkan sisi dadu 4 adalah $\frac{5}{6}$

Frekuensi harapan suatu kejadian Frekuensi harapan suatu kejadian adalah hasil kali munculnya suatu kejadian dengan banyaknya percobaan yang dilakukan

$$Fh = P\ (A) \times n$$

Contoh:

Pada pelemparan sebuah koin, nilai peluang munculnya gambar adalah $\frac{1}{2}$ apabila pelemparan koin dilakukan sebanyak 30 kali maka berapkah harapan munculnya gambar?

Jawaban:

$Fh = P\ (A) \times n$

$Fh = 30$

$Fh = 15$ kali

Jadi harapan munculnya gambar dari 30 kali pelemparan dadu adalah 15 kali.

9.4.2 Kejadian Tidak Saling Lepas

Suatu peristiwa atau kejadian yang tidak saling lepas (*non mutually exclusive)* terjadi apabila kejadian satu dapat terjadi secara bersama-sama dengan kejadian lain.

Suatu operasi gabungan dari elemen (n) himpunan E dan elemen (n) himpunan F yang tidak saling lepas dapat dirumuskan:

$$(E \cup F) = n(E) + n(F) - n(E \cap F) \tag{9.6}$$

Probabilitas gabungan dua kejadian E dan kejadian F yang tidak saling lepas dalam ruang contoh S dapat ditentukan dengan membagi persamaan (9.6) dengan n(S):

$$\frac{n\ (E \cup F)}{n(S)} = \frac{n(E)}{n(S)} + \frac{n(F)}{n(S)} - \frac{n(E \cap F)}{n(S)} \tag{9.7}$$

Berdasarkan persamaan (9.1), maka:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F) \tag{9.8}$$

Sebaliknya:

$$P(E \cap F) = P(E) + P(F) - P(E \cup F) \tag{9.9}$$

Contoh:

Terdapat suatu acara seminar nasional yang melibatkan 300 peserta. Acara ini diikuti oleh sebanyak 180 Perempuan (P), 150 peserta dari kalangan pekerja swasta (S), 90 peserta dari kalangan perempuan pegawai swasta. Jika seorang peserta seminar nasioanl mendapatkan hadiah undian, berapakah probabilitas peserta yang terpilih: a.) bukan pegawai swasta , b.) perempuan atau pegawai swasta, c.) perempuan tetapi bukan pegawai swasta.

Probabilitas Perempuan $P(P): \frac{180}{300} = 0,6$

Probabilitas pegawai swasta $P(S): \frac{150}{300} = 0,5$

Probabilitas perempuan pegawai swasta $P(P \cap S) = \frac{90}{300} = 0,3$

a.) Probabilitas bukan pegawai swasta:

$$P(S') = 1 - P(S) = 1 - 0,5 = 0,5$$

b.) Probabilitas perempuan atau pegawai swasta:

$$P(P \cup S) = P(P) + P(S) - P(P \cap S) = 0,6 + 0,5 - 0,3 = 0,8$$

c.) Probabilitas perempuan tetapi bukan pegawai swasta:

$$P(P \cap S') = \frac{(180 - 150)}{300} = \frac{30}{300} = 0,1$$

Contoh:

Seorang mahasiswa memiliki kemungkinan lulus dalam mata kuliah statistik sebesar 0,80; selanjutnya lulus mata kuliah ekonometrik sebesar 0.75 dan lulus keduanya 0,70. Hitunglah: a.) Probabilitas mahasiswa tersebut akan lulus dalam matakuliah statistik dan ekonometrik b.) Mahasiwa tidak lulus keduanya.

a.) Probabilitas seorang mahasiswa lulus kedua mata kuliah:

P(S) = 0,80

P(E) = 0,75

P(S ∩ E) = 0,70

$$\text{Jadi, } P(S \cup E) = P(S) + P(E) - P(S \cap E)$$
$$= 0{,}80 + 0{,}75 - 0{,}70$$
$$= 0{,}85$$

b.) Probabilitas seorang mahasiswa yang tidak lulus kedua mata kuliah:

$$\mathrm{P}(S') \cap \mathrm{P}(\mathrm{E}') = \mathrm{P}(\mathrm{S} \cup \mathrm{E})'$$
$$= 1 - P(S \cup E) = 1 - 0{,}85 = 0{,}15$$

9.5 KEJADIAN TIDAK BEBAS

Suatu peristiwa atau dua peristiwa atauu lebih bisa dikatakan tidak bebas (*dependen)* jika suatu peristiwa atau kejadian satu dipengaruhi oleh kejadian yang lain. Probabilitas tidak bebas dibedakan menjadi tiga: probabilitas marginal (*marginal probability*), probabilitas bersama (*joint probability*) dan probabilitas bersyarat (*conditional probability*).

9.5.1 Probabilitas Marginal

Probabilitas marginal atau probabilitas tunggal adalah probabilitas terjadinya satu kejadian. Secara matematis, probabilitas marginal dapat dinyatakan P(E) = probabilitas terjadinya peristiwa E. jika kejadian E dan F merupakan kejadian yang tidak bebas atau E terjadi terlebih dahulu maka P(F|E) merupakan probabilitas kejadian F dengan syarat kejadian E terjadi lebih dahulu.

9.5.2 Probabilitas Bersama

Probabilitas yang terjadi saat dua peristiwa atau lebih terjadi pada saat yang sama. Dalam hukum penjumlahan, jika kejadian E dan F merupakan kejadian yang bersifat *non mutually exclusive*, maka $P(E \cap F) = P(E) + P(F) - P(E \cup F)$ dengan $P(E \cap F) \neq 0$. Dan dalam hukum perkalian, suatu kejadian E dan F merupakan dua kejadian yang tidak bebas secara statistik, sehingga berdasarkan hukum perkalian, maka $P(E \cap F) = P(E) \,.\, P(E|F)$.

Contoh:
Sebuah Survei Ekstrakulikuler sekolah telah dilakukan, dari 250 siswa sekolah. Dari hasil survey tersebut, diperoleh hasil bahwa 150 siswa memilih ektrakulikuler pramuka, sedangkan 175 siswa memilih ekstrakulikuler PMR. Berapakah probabilitas bahwa seorang siswa telah terpilih ikut dalam ekstrakulikuler Pramuka atau PMR?

Menerapkan aturan penjumlahan khusus yaitu *mutually exclusive* atau kejadian saling lepas, maka probabilitas seorang siswa terpilih mengikuti kegiatan ekstrakulikuler Pramuka adalah 0,6 yang diperoleh dari (150/250). Dan probabilitas seorang siswa terpilih mengikuti kegiatan ekstrakulikuler PMR adalah 0,7 yang diperoleh dari (175/250). Jumlah probabilitas kedua kegiatan ini akan lebih dari 1 (0,6 + 0,7 = 1,3). Namun setelah disurvey kembali, terdapat 50 siswa yang telah memilih kedua kegiatan ekstrakulikuler. Maka dari itu, probabilitas siswa memilih ekstrakulikuler Pramuka dan PMR adalah:

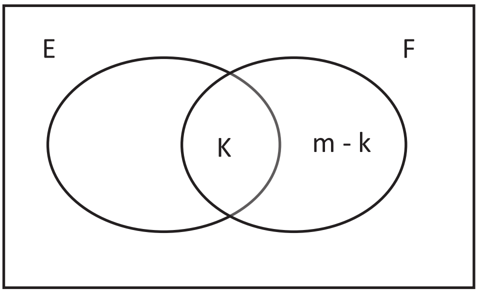
$$P(\text{Pramuka} \cap \text{PMR}) = P(\text{Pramuka}) + P(\text{PMR}) - P(\text{Pramuka} \cup \text{PMR})$$

$$\frac{150}{250} + \frac{175}{250} - \frac{50}{250} = 0{,}6 + 0{,}7 - 0{,}2 = 1{,}1$$

Bisa disimpulkan probabilitas siswa memilih kedua kegiatan ektrakulikuler adalah 1,1.

9.5.3 Probabilitas Bersyarat

Probabilitas terjadinya E merupakan sebuah kejadian dalam ruang contoh S bila diketahui kejadian lain F telah terjadi disebut Probabilitas bersyarat (*conditional probability*) yang dinotikasikan dalam bentuk P(E|F) atau bisa dikatakan bahwa probabilitas terjadinya E bila kejadian F telah terjadi. Suatu kejadian dapat bergantung pada terjadi atau tidaknya suau kejadian lain. Untuk kejadian yang bergantung pada kejadian lain, nilai peluangnya dicari dengan menggunakan peluang bersyarat



GAMBAR 9.2 ***Probabilitas Bersyarat***

Diasumsikan bahwa ruang contoh S mempunyai *n* kejadian. Kejadian E mempunyai *m* elemen dan kejadian (E ∩ F) mempunyai *k* elemen (k ≤ m). dengan menggunakan prinsip umum probabilitas,

$$P(F) = \frac{m}{n}\ dan\ P(E \cap F) = \frac{k}{n} \tag{9.10}$$

Menetukan probabilitas bersyarat kejadian E jika kejadian F telah diketahui yaitu P(E|F). saat F telah terjadi, hanya *m* elemen yang ada didalam kejadian F, dan terdapat *k* elemen terdapat di kejadian E karena P(E ∩ F) mempunyai *k* elemen. Maka bisa di tulis rumusnya:

$$P(E|F) = \frac{k}{m} \tag{9.11}$$

Dengan membagi pembilang dan penyebut dengan *n* maka:

$$P(E|F) = \frac{k/n}{m/n} \tag{9.12}$$

Substitusi bentuk ini pada persamaan (9.11) sehingga:

$$P\ (E|F) = \frac{P(E \cap F)}{P(F)} = untuk\ P(F) > 0 \tag{9.13}$$

Contoh:
Probabilitas perjalanan Kereta api Argo Parahyangan berangkat tepat waktu (PW) = 0,95. Perjalanan kereta api argo parahyangan smapai menuju ke tempat tujuan sesuai dengan jadwal P(J) = 0,80 dan probabilitas perjalanan kereta api berangkat tepat waktu dan sampai tujuan sesuai jadwal adalah P(W∩J) = 0,60.

a.) Berapakah probabilitas kereta berangkat tepat waktu bila diketahui bahwa kereta api argo parahyangan sampai tujuan sesuai jadwal?

$$P(W|J) = \frac{P(W \cap J)}{P(J)} = \frac{0,60}{0,80} = 0,75$$

b.) Berapakah probabilitas kereta api argo parahyangan sampai tujuan sesuai jadwal jika diketahui kereta api berangkat tepat waktu?

$$P(J|W) = \frac{P(J \cap W)}{P(W)} = \frac{0,60}{0,95} = 0,63$$

9.6 KEJADIAN BEBAS

Dua peristiwa atau lebih dikatakan bebas (*independent*) jika, kejadian yang satu tidak dipengaruhi kejadian yang lain. Probabilitas dari kejadian bebas ini dapat dibedakan menjadi tiga yaitu: probabilitas marginal, probabilitas Bersama dan probabilitas bersyarat.

9.6.1 Probabilitas Marginal

Probabilitas marginal adalah probabillitas yang terjadi saat satu kejadian yang tidak memiliki hubungan dengan kejadian yang lain. Secara matematis, probabilitas marginal dirumuskan sebagai P(E) = probabilitas terjadinya peristiwa. Probabilitas terjadinya peristiwa F dituliskan sebagai P(F). namun, kedua peristiwa ini tidak memiliki hubungan dan tidak saling mempengaruhi kejadian yang lainnya. Kejadian ini bersifat bebas secara statistik.

9.6.2 Probabilitas Bersama

Probabilitas Bersama adalah probabilitas dua peristiwa atau lebih yang terjadi secara Bersama-sama atau secara berurutan. Probabilitas Bersama merupakan hasil perkalian dari probabilitas marginal atau probabilitas masing-masing peristiwa. Dalam hukum perkalian, suatu kejadian E dan F merupakan dua kejadian bebas. Jadi, tidak atau terjadinya kejadian E, probabilitas peristiwa F akan sama saja.

Misalkan $P(E) = \frac{n\,(E)}{n(S)}\; dan\; P\;(F) = \frac{n(F)}{n(S)}$, Bila seluruh anggota himpunan ruang contoh kejadian (E∩F) ditulis n(E∩F) = n(E) . n(F), maka probabilitas bersama kejadian bebas adalah:

$$P(E \cap F) = \frac{n(E)}{n(S)}\; x\; \frac{n(F)}{n(S)} = P(E).P(F) \qquad (9.14)$$

P(E∩F) adalah probabilitas terjadinya peristiwa E dan peristiwa F bersama atau berurutan. P(E) dan P(F) masing-masing merupakan probabilitas marginalnya.

Contoh:
Probabilitas munculnya Gambar koin A adalah P(A) = 0,8 dan angka P(S) = 0,6. Jika dilakukan 3 kali pelemparain uang koin, berapa probabilitas a.) Munculnya sisi gambar A-A-S. b.) Gambar A-A-A dan c.) muncul S dalam 3 kali lemparan.

a.) Muncul sisi gambar koin A-A-S:

$$P(\,A_1\, \cap\, S\, \cap A_2) = P(A_1\,).P(S)\,.P(A_2\,)$$
$$= 0{,}8\; x\; 0{,}6\; x\; 0{,}8 = 0{,}384$$

b.) Munculnya sisi gambar A-A-A:

$$P(A_1\, \cap A_2\, \cap A_3\,) = \; P(\,A_1\,)\,.P(A_2\,).P(A_3\,)$$
$$= \; 0{,}8\; x\; 0{,}8\; x\; 0{,}8 = 0{,}512$$

c.) Probabilitas munculnya S dalam 3 kali pelemparan:

$$P(S) = P(S\, \cap A_1\, \cap A_2\,) + P(\,A_1\, \cap\, S \cap A_2\,) + P(A_1\, \cap A_2\, \cap S)$$

= 0,384 + 0,384 + 0,384

= 1,152

9.6.3 Probabilitas Bersyarat

Suatu kejadian E dan F dikatakan bersifat bebas secara statistik jika terjadinya peristiwa yang satu tidak akan mempengaruhi peristiwa yang lain. Terjadi tidaknya peristiwa E, popularitas kejadian F akan sama saja sehingga

P(E|F) = P(E). implikasi dari probabilitas yang bersifat bebas ini mengharuskan P(F|E) = P(F) dan sebaliknya. Oleh karena itu, P(E∩F) setara dengan P(F∩E).

$$P(E \cap F) = P(E|F)x\, P(F) \quad (9.15)$$

Oleh karena P(E|F) = P(E) maka:

$$P(E \cap F) = P(E).P(F) \qquad (9.16)$$

Karena dalam peristiwa yang bersifat bebas bentuk persyaratan tidak berpengaruh, maka dapat dinyatakan bahwa P(F|E) = P(F) dengan P (F∩E) = P(F) . P(E). Perluasan probabilitas suatu kejadian-kejadian E_1, E_2, E_3, …E_n yang bersifat bebas menghasilkan bentuk umum:

$$P(E_1 \cap E_2 \cap E_3 \cap \ldots . \cap E_n) = P(E_1).P(E_2)$$
$$.P(E_3)\ldots .P(E_n) \qquad (9.17)$$

Contoh:
Dalam perlombaan memanah, dilaporkan bahwa H telah 6 kali mengenai sasaran dalam 10 sasaran panah. Sedangkan A sebanyak 4 kali dalam 8 sasaran panah. Berapakah probabilitas jika dalam sasarannya:

a.) H tidak mengenai sasaran memanah
b.) A tidak mengenai sasaran memanah
c.) H dan A semuanya mengenai sasaran
d.) H dan A semuanya tidak mengenai sasaran

P(H) = probabilitas memanah H mengenai sasaran
P(H′) = probabilitas memanah H tidak mengenai sasaran
P(A) = probabilitas memanah A mengenai sasaran
P(A′) = probabilitas memanah A tidak mengenai sasaran

Maka:

a.) $P(H) = \frac{6}{10} \rightarrow P(H') = 1 - P(H) = 1 - \frac{6}{10} = \frac{8}{20}$

Probabilitas H tidak mengenai sasaran 0,4.

b.) $P(A) = \frac{4}{8} \rightarrow P(A') = 1 - P(A) = 1 - \frac{4}{8} = \frac{10}{20}$

Probabilitas A tidak mengenai sasaran 0,5

c.) H dan A semuanya mengenai sasaran:

$P(H \cap A) = P(H).P(A) = \frac{6}{10} . \frac{4}{8} = \frac{6}{20}$

d.) H dan A semuanya tidak mengenai sasaran:

$$P(H' \cap A') = P(H') . P(A') = \frac{2}{5} . \frac{2}{4} = \frac{4}{20}$$

Contoh:

Sebuah dadu dilempar sekali tentukan peluang munculnya mata dadu genap dengan syarat munculnya kejadian mata dadu prima terlebih dahulu.

Jawaban

Misal A adalah kejadian munculnya mata dadu prima

Ruang sampel: s = {1,2,3,4,5,6}, sehinga $n(s) = 6$

A = {2,3,5}, sehingga $n(A) = 3$

Peluang kejadian A: $P(A) = \frac{3}{6} = \frac{1}{2}$

Misal B adalah kejadian munculnnya mata dadu genap

B = {2,4,6), sehingga irisannya $A \cap B = \{2\}$, dengan $n(A \cap B) = 1$

Peluang kejadian $P(A \cap B) = \frac{n(A \cap B)}{n(s)} = \frac{1}{6}$

Jadi, peluang munculnya mata dadu genap dengan syarat munculnya kejadian mata dadu prima lebih dahulu:

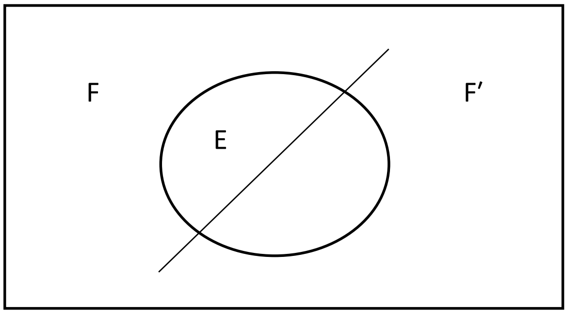
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

$$P(B \mid A) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

Jadi, peluang munculnya mata dadu genap dengan syarat munculnya kejadian mata dadu prima lebih dahulu adalah $\frac{1}{3}$

9.7 TEORI BAYES

Teori Bayes sangat penting dalam penerapan probabilitas bersyarat. Teori ini pertama kali dikenalkan oleh Thomas Bayes tahun 1702-1763. Teori bayes merupakan teori yang memperbaiki atau merevisi suatu probabilitas dengan cara memanfaatkan informasi tambahan. Dari probabilitas awal (*prior probability)* yang belum diperbaiki dirumuskan berdasarkan informasi yang tersedia kemudian dibentuklah probabilitas berikutnya (*posterior probability*).



GAMBAR 9.3 ***Diagram Venn yang menunjukkan kejadian E, F dan F′***

Berdasarkan diagram venn diatas, dapat disimpulkan E sebagai paduan dua kejadian yang saling terpisah E ∩ F dan E ∩ F′ sehingga:

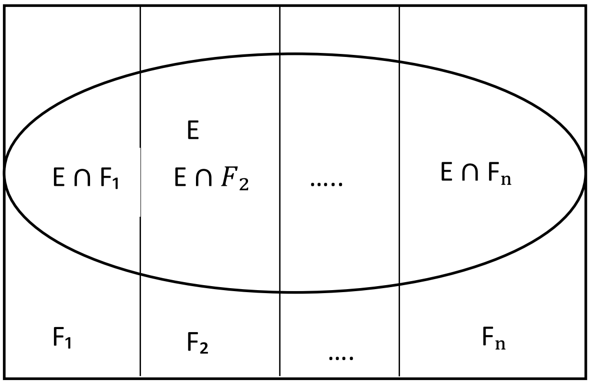
$$E = (E \cap F) \cup (E \cap F') \qquad (9.18)$$

Probabilitas kejadian E dapat dituliskan:

$$\begin{aligned} P(E) &= P(E \cap F) + P(E \cap F') \\ &= P(F)\,.\,P(E|F) + P(F')\,.\,P(E|F') \quad (9.19) \end{aligned}$$

Selain untuk menghitung P(E), pertanyaan selanjutnya adalah menentukan berapa besar probabilitas bersyarat $P(F_i|E)$ seperti pada gambar 9.4.pertanyaan ini dapat dijelaskan dengan menggunakan teori bayes yaitu jika kejadian-kejadian F_1, F_2, …, F_n merupakan penyekatan dari ruang sampel dengan $P(F_i) \neq 0$ untuk i = 1,2,3,…n.

$$P(F_i\,|E) = \frac{P(F_i \cap E)}{P(E)} \quad \text{dengan} \quad P(E) > 0 \qquad (9.20)$$



GAMBAR 9.4 *Diagram Penyekatan Ruang*

Jika kejadian-kejadian $F_1, F_2, \ldots, F_n$ adalah kejadian bebas satu sama lain yang juga merupakan kejadian kolektif dan terbatas dengan probabilitas priori $P(F_1)$, $P(F_2)$ … $P(F_n)$ serta karena kejadian E dibatasi kejadian F_1, F_2, … dan F_n maka kejadian E merupakan gabungan dari kejadian-kejadian $(E \cap F_1)$, $(E \cap F_2)$, ..., $(E \cap F_n)$. Maka probabilitas P(E) ditulis sebagai:

$$P(E) = P(E \cap F_1) + P(E \cap F_2) + \cdots + P(E \cap F_n) \qquad (9.21)$$

Oleh karena itu, kesamaan probabilitas $P(E \cap F_i) = P(F_i \cap E)$, memungkinkan bahwa:

$$P(F_i \cap E) = P(F_i) \,.\, P(E|F_i) \qquad (9.22)$$

Dengan mensubstitusikan persamaan (9.22) dan (9.21) diperoleh:

$$(E) = P(F_i) \,.\, P(E|F_1) + P(F_2) \,.\, P(E|F_2) + \cdots + P(F_n) \,.\, P(E|F_n) \qquad (9.23)$$

Selanjutnya dengan mengeliminasi penyebut dan pembilang pada (9.19) dengan dua persamaan yang terakhir, selanjutnya probabilitas bersyarat $P(F_iE)$ untuk sembarang kejadian E:

$$P(F_i|E) = \frac{P(F_i) \,.\, P(E|F_i)}{P(F_1) \,.\, P(E|F_1) + P(F_2) \,.\, P(E|F_2) + \cdots + P(F_n) \,.\, P(E|F_n)} \qquad (9.24)$$

Bentuk rumus diatas merupakan penyelesaian umum dari Teori Bayes.

Contoh:

PT. CARITA memproduksi produk makanan dengan menggunakan 3 mesin yaitu mesin A, mesin B, mesin C. Mesin A mampu memproduksi 900 unit, mesin B mampu memproduksi 450 unit, dan mesin C mampu memproduksi 150 unit. Namun terdapat masalah dalam proses produksi. Tercatat sebanyak 15% produk makanan rusak yang diproduksi mesin A, sebanyak 10% dari mesin B, dan sebanyak 5% dari mesin C. jika dari jumlah produksi dipilih 1 produk makanan secara acak, hitunglah:

a.) Probabilitas produk yang dipilih adalah rusak P(E)
b.) Probabilitas produk yang dipilih baik P(E′)
c.) Bila produk yang dipilih rusak, maka berapa probabilitasnya yang diproses mesin A

Jawab:

Pertama, dihitunglah probabilitas produk makanan yang dihasilkan oleh masing-masing mesin:

a.) a. Probabilitas produk makanan pada mesin $P(A) = \frac{900}{1500} = 0{,}6$

Probabilitas produk makanan rusak dari mesin A, $P(E|A) = 0{,}15$

b. Probabilitas produk makanan pada mesin $P(B) = \frac{450}{1500} = 0{,}3$

Probabilitas produk makanan rusak dari mesin B, $P(E|B) = 0{,}10$

c. Probabilitas produk makanan pada mesin $P(C) = \frac{150}{1500} = 0{,}1$

Probabilitas produk makanan rusak dari mesin C, $P(E|C) = 0{,}05$

Jadi, Probabilitas produk yang dipilih rusak P(E) adalah:

$$P(E) = P(A)\,.\,P(E|A) + P(B)\,.\,P(E|B) + P(C)\,.\,P(E|C)$$
$$= 0{,}6(0{,}15) + 0{,}3(0{,}10) + 0{,}1(0{,}05)$$
$$= 0{,}125$$

b.) $P(A) = 0{,}6 \;\; P(E'|A) = 1 - P(E|A) = 0{,}85$

$P(B) = 0{,}3 \;\; P(E'|B) = 1 - P(E|B) = 0{,}9$

$P(C) = 0{,}1 \; P(E'|C) = 1 - P(E|C) = 0{,}95$

Jadi, Probabilitas produk makanan yang dipilih baik P(E′) adalah

$$P(E') = P(A)\,.\,P(E'|A) + P(B)\,.\,P(E'|B) + P(C)\,.\,P(E'|C)$$
$$= 0{,}6(0{,}85) + 0{,}3(0{,}9) + 0{,}1(0{,}95)$$
$$= 0{,}875$$

c.) Probabilitas produk makanan dari mesin A diketahui ternyata rusak adalah:

$$P(A|E) = \frac{(0{,}6)\,.\,(0{,}15)}{(0{,}6)(0{,}15)+(0{,}3)(0{,}10)+(0{,}1)(0{,}05)}$$
$$= \frac{0{,}09}{0{,}125} = 0{,}72$$

Bab 10
Distribusi Peluang

10.1 PENGANTAR

Jika pada bab sebelumnya kita telah membahas tentang peluang, maka pada bab ini akan diulas lebih dalam apa itu distribusi sampling. Secara definisi distribusi sampling adalah salah satu komponen dalam ilmu statistika yang berfungsi untuk filterisasi data yang mempunyai masalah perhitungan peluang berdasarkan jenis peubah acaknya, baik diskrit atau kontinu. Sehingga, fokus pembahasan dalam bab ini adalah diharapkan dapat membedakan peubah acak diskrit dan kontinu; peluang kejadian distribusi binomial, poisson dan normal; dan menentukan peluang seuatu kejadian dengan menggunakan distribusi Binomial melalui pendekatan distribusi Poisson dan Normal.

10.2 DEFINISI DISTRIBUSI PELUANG

Distribusi peluang merupakan alat bagi seorang peneliti untuk menentukan apa yang dapat peneliti harapkan, apabila asumsi-asumsi yang dibuat oleh peneliti benar. Distribusi peluang memungkinkan para pembuat keputusan untuk memperoleh dasar logika yang kuat di dalam keputusan, dan sangat berguna sebagai dasar pembuatan ramalan berdasarkan informasi yang terbatas atau pertimbangan-pertimbangan teoritis, serta berguna pula untuk menghitung probabilitas terjadinya suatu kejadian. Setiap kejadian yang dapat dinyatakan sebagai perubahan nilai suatu peubah, umumnya mengikuti suatu distribusi peluang tertentu dan apabila sudah diketahui jenis distribusinya, maka peneliti akan dengan mudah dapat mengetahui besarnya nilai probabilitas terjadinya kejadian tersebut. Misalnya bila dilakukan undian menggunakan mata uang logam lima ratus rupiah, akan diperoleh kemungkinan hasil:

Sebagai contoh, ruang sampel dari suatu eksperimen melempar satu mata uang tiga kali (Muka disimbolkan M dan belakang disimbolkan B) adalah

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

Apabila konsentrasi kita terpusat bagian muka (M) maka hasil yang mungkin terjadi dari kemunculan dalam tiga kali lemparan adalah 0,1,2,3. Bilangan 0,1,2,3 ini disebut dengan jumlah pengamatan acak dari kemunculan bagian muka. Kemunculan kejadian ini seterusnya akan disebut dengan peubah acak X, yaitu nilai yang menyatakan banyaknya kejadian munculnya bagian muka (M). Secara umum, peubah acak dibagi menjadi dua:

1. Peubah acak diskrit, yakni apabila nilai dari peubah acak adalah bilangan bulat atau jika banyaknya titik sampel dari suatu ruang sampe berhingga banyaknya. Contohnya adalah jumlah pekerja yang harus di PHK, banyaknya jumlah rumah yang sedang dalam proses pembangunan
2. Peubah acak kontinu, yakni apabila nilai dari peubah acak adalah dalam bentuk pecahan, bilangan desimal atau bilangan riil. Sebagai contoh, berat badan seseorang atau tinggi badan.

Contoh:
P(muka G): P(muka M) = 1/2
Bila dipandang muka G yang nampak, maka muka M = 0 G, dan muka G = 1G.
Bila banyak muka G yang muncul diberi simbol, maka bila muncul M berlaku X = 0, dan bila muncul muka G, X = 1, atau dapat ditulis:

$P(X = 0) = 1/2$

$P(X = 1) = 1/2$

Lebih lanjut kalau eksperimen dikerjakan dengan dua mata uang lima ratus rupiah, maka peristiwa yang terjadi adalah: GG, GM, MG, dan RR. Ditulis P(GG), P(GM), P(MG), dan P(RR). Bila X menyatakan banyaknya kali muka G, maka:

X = 0, untuk (RR) diperoleh P(X = 0) = 1/4.

X = 1, untuk (GR) diperoleh P(X = 1) = 1/4 + 1/4 = 1/2.

X = 0, untuk (GG) diperoleh P(X = 2) = 1/4.

Berdasarkan harga-harga tersebut, dapat ditampilkan dalam bentuk tabel, seperti yang ditampilkan pada Tabel 2.6.

TABEL 10.1 *Harga-harga P(X)*

X	P(X)
0	¼
1	½
2	¼
Jumlah	1

Bila dilanjutkan dengan tiga mata uang lima ratus rupiah, maka akan diperoleh delapan periswa, yaitu GGG, GGR, GRG, RGG, RRG, RGR, GRR, dan RRR. Peluang untuk masing-masing peristiwa adalah 1/8 maka dari X = 0, 1, 2, 3. Diperoleh harga-harga P(X = 0) = 1/8; P(X = 1) = 3/8; P(X = 2) = 3/8; dan P(X = 3) = 1/8. Jika hasil disusun dalam sebuah tabel, maka dapat dilihat pada Tabel 2.7.

TABEL 10.2 *Harga-harga P(X)*

X	P(X)
0	1/8
1	3/8
2	3/8
3	1/8
Jumlah	1

Proses ini dapat diteruskan untuk undian dengan menggunakan empat mata uang, lima mata uang, dan seterusnya. Notasi X di atas hanya memiliki harga 0, 1, 2, 3, dan seterusnya, disebut peubah acak diskrit. Untuk peubah acak dapat ditentukan nilai ekspektasinya, yaitu:

$$E(X) = \Sigma Xi \,.\, P(Xi) \qquad \text{(Rumus 2.8)}$$

Misalnya pengamatan memerhatikan bahwa banyak kendaraan yang melalui sebuah tikungan setiap menit mengikuti distribusi peluang berikut:

TABEL 10.3 ***Hasil Pengamatan Banyak Mobil yang Lewat***

Jumlah Kendaraan	0	1	2	3	4	5	6	7	8
Peluang	0,01	0,05	0,1	0,28	0,22	0,18	0,08	0,05	0,03

Berdasarkan Tabel di atas mobil yang lewat melalui tikungan itu sebanyak:

$$= 1 – (0{,}01 + 0{,}05 + 0{,}1) = 0{,}84.$$

Selanjutnya dengan rumus: $E(X) = \Sigma Xi \,.\, P(Xi)$, diperoleh:

$$\begin{aligned} E(X) &= (0)(0{,}01) + (1)(0{,}05) + (2)(0{,}1) + (3)(0{,}28) + (4)(0{,}22) \\ &\quad + (5)(0{,}18) + (6)(0{,}08) + (7)(0{,}05) + (8)(0{,}03) = 3{,}94 \end{aligned}$$

Sehingga dapat ditarik simpulan bahwa terdapat 394 mobil yang lewat setiap 100 menit. Peubah acak yang diskrit disebut juga peubah acak kontinu. Salah satu ciri dari peubah acak kontinu adalah bahwa peristiwa peubah ini memiliki harga sebarang, dapat pula berupa pecahan atau bentuk desimal. Berikut ini akan diuraikan macam-macam distribusi, yaitu distribusi binomial, distribusi multinomial, distribusi poisson, distribusi hipergeometrik, distribusi normal, distribusi student, distribusi *Chi Square*, dan distribusi fisher (F).

10.3 DISTRIBUSI PELUANG DISKRIT

10.3.1 Distribusi Binomial

Distribusi binomial adalah suatu distribusi peluang yang menggunakan peubah acak diskrit, yang terdiri dari dua kejadian yang berkomplemen, seperti sukses-gagal, baik-cacat. Suatu eksperimen yang menghasilkan dua kemungkinan peristiwa A dan bukan A (atau A') yang bersifat dikotomi dengan P(A) = π serta memiliki harga tetap untuk peristiwa A. Jika eksperimen tersebut diulang-ulang akan didapatkan distribusi Bernoulli. Lazimnya suatu eksperimen dapat dikatakan sebagai eksperimen binomial, apabila memenuhi syarat, yaitu: (1) setiap percobaan menghasilkan dua kejadian, seperti kelahiran anak yakni laki-laki – perempuan, transaksi saham jual-beli, atau perkembangan suku bunga yakni naik-turun; (2) setiap eksperimen mempunyai dua hasil yang dikatagorikan menjadi sukses dan gagal; (3) probabilitas suatu kejadian untuk suskes atau gagal adalah tetap untuk setiap kejadian. Misalnya P(p), peluang sukses, P(q) peluang gagal, dan P(p) + P(q) = 1; (4) probabilitas sukses sama pada setiap eksperimen; (5) eksperimen tersebut harus bebas satu sama lain, artinya hasil eksperimen yang satu tidak mempengaruhi hasil eksperimen lainnya; dan (6) data yang dihasilkan adalah data perhitungan.

Percobaan Bernoulli sebanyak N kali secara independen, sehingga X diantaranya menghasilkan peristiwa A dan sisanya (N-X) peristiwa A'. Jika π = P(A) untuk tiap percabaan, maka 1 – π = P(A'), sehingga peluang terjadinya A sebanyak X = x kali diantara N, yaitu:

P(X) = P(X = x) = $(Px\ N)$ πx
(1 – π) (1 – π)N-x (Rumus 2.9)

Dimana:

Nilai = 0, 1, 2, ... , N; 0 < π < 1

$$P_x^N = \frac{N!}{x!(N-x)!}$$

Hubungan yang dinyatakan dalam Rumus 2.8 merupakan distribusi dengan peubah acak diskrit dan dinamakan **distribusi binomial**, dengan koefisien binomial pada Rumus 2.9. Distribusi binomial memiliki dua parameter μ dan σ yang persamaannya adalah:

$$Nilai\ Rata-Rata: \qquad \mu_x = n\,.\,p_x$$

$$Variasni: \qquad \sigma_x^2 = n\,.\,p_x\,.\,(1 - p_x)$$

Misalnya menghitung peluang untuk memeroleh 6 kali muka G (mata uang logam lima ratus rupiah), bila dilakukan 10 lemparan.

$P(X = 6) = (P^{10}_{6})\ 1/2^6(1 - 1/2)^{10-6}$
$= 0{,}205$

Contoh:
Suatu eksperimen binomial, yang terdiri dari pengambilan satu bola secara acak dari kotak yang berisi 30 bola merah (30M) dan 70 bola putih (70P). Y adalah peubah acak dengan nilai sebagai berikut.

Y = 0, kalau bola putih yang terambil
Y = 1, kalau bola merah yang terambil

P(M) = p = probabilitas untuk mendapat bola merah (sukses)
= 30/100
= 0,30

P(P) = q = probabilitas untuk mendapat bola putih (gagal)
= 70/100
= 0,70

E(Y) = 1(p) + 0(q)
= 1(0,3) + 0(0,7)
= 0,3

Bila dilakukan eksperimen empat kali. Pengambilan bola dilakukan dengan pengembalian bola yang terambil. Hal ini untuk menjaga agar eksperimen yang satu tidak mempengaruhi hasil eksperimen yang lain. Eksperimen ini akan menghasilkan 24 = 16 hasil, yakni:

MMMM	MPMM	PMMM	PPMM
MMMP	MPMP	PMMP	PPMP
MMPM	MPPM	PMPM	PPPM
MMPP	MPPP	PMPP	PPPP

Setiap hasil eksperimen terdiri dari empat kejadian yang bebas satu sama lain, sehingga probabilitas terjadinya setiap hasil eksperimen merupakan hasil kali probabilitas masing-masing kejadian, misalnya P(MMPM) = ppqp = (0,3)(0,3)(0,7)(0,3) = 0,0189. Aturan perkalian untuk kejadian-kejadian bebas dan aturan penjumlahan untuk kejadian-kejadian yang saling meniadakan, yang sudah dibahas sebelumnya dapat diterapkan di sini dan perhitungannya ialah:

$$\begin{aligned} P(3M \text{ dan } 1P) &= P(MMMP) + P(MMPM) + P(MPMM) + P(PMMM) \\ &= ((0,3)\ (0,3)\ (0,3)\ (0,7)) + ((0,3)\ (0,3)\ (0,7)\ (0,3)) \\ &\quad + ((0,3)\ (0,7)\ (0,3)\ (0,3)) \\ &= + ((0,7)\ (0,3)\ (0,3)\ (0,3)) \\ &= 0,0756 \end{aligned}$$

Tanpa memperhatikan urutan dari masing-masing kejadian, setiap suku dalam penjumlahan tersebut mempunyai probabilitas sebesar pppq = p3q. Dengan cara yang sederhana ini, dapat menghitung probabilitas untuk mendapatkan sejumlah bola merah tertentu sebagai hasil eksperimen. Dapat ditunjukkan bahwa apabila eksperimen dilakukan sebanyak 4 kali, maka X = 0, 1, 2, 3, 4. Sedangkan untuk n kali, ialah X = 0, 1, 2, … , n. Apabila semua nilai probabilitas X sebagai hasil suatu eksperimen dihitung, akan diperoleh distribusi probabilitas X dan disebut distribusi probabilitas binomial.

$$P(X = 0) = P(PPPP) = (0,7)(0,7)(0,7)(0,7) = (0,7)^4 = 0,2401$$

$P(X = 1) = pq^3 + qpq^2 + q^2pq + q^3p$

$= (0{,}3)(0{,}7)^3 + (0{,}7)(0{,}3)(0{,}7)^2 + (0{,}7)^2\,(0{,}3)(0{,}7) + (0{,}7)^3\,(0{,}3)$

$= 0{,}4116$

$P(X = 2) = p^2q^2 + pqpq + pq^2p + qp^2q + qpqp + q^2p^2$

$= (0{,}3)^2(0{,}7)^2 + (0{,}3)(0{,}7)(0{,}3)(0{,}7) + (0{,}3)(0{,}7)^2(0{,}3)$

$+ (0{,}7)(0{,}3)^2(0{,}7)$

$= + (0{,}7)(0{,}3)(0{,}7)(0{,}3) + (0{,}7)^2(0{,}3)^2$

$= 0{,}2646$

$P(X = 3) = p^3q + p^2qp + pqp^2 + qp^3$

$= (03)^3(0{,}7) + (0{,}3)^2(0{,}7)(0{,}3) + (0{,}3)(0{,}7)(0{,}3)^2 + (0{,}7)(0{,}3)^3$

$= 0{,}0756$

$P(X = 4) = P(MMMM) = p^4 = (0{,}3)^4$

$= 0{,}0081$

Berdasarkan contoh soal di atas, dapat disimpulkan bahwa dalam distribusi probabilitas binomial, dengan n percobaan, berlaku rumus:

Pr(x sukses, dalam n percobaan) = p^xq^{n-x}

Dimana:

x = 0, 1, 2, 3, …, n

p = probabilitas sukses

q = (1 – p) = probabilitas gagal

Aturan umum permutasi dapat digunakan untuk memperoleh banyaknya kemungkinan urutan yang berbeda, dimana masing-masing urutan terdapat x sukses, misalnya x = 3 (3 sukses), sehingga: MMMP, MMPM, MPMM, PMMM. Apabila suatu himpunan yang terdiri dari n elemen dibagi dua, yaitu x sukses dan (n – x) gagal, maka banyaknya permutasi dari n elemen yang diambil x setiap kali dapat dihitung berdasarkan rumus kombinasi:

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

Disebut koefisien binomial (merupakan kombinasi dari n elemen yang diambil x setiap kali). Masing-masing probabilitas pada distribusi binomial dihitung dengan rumus:

$$p_r(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

Dimana

$x = 0, 1, 2, \ldots, n$.

$p_r(x)$ dari rumus tersebut merupakan fungsi probabilitas, karena:

$$\frac{n!}{x!(n-x)!}$$

a. $p_r(x) \geq 0$, untuk semua x, sebab ≥ 0 dan $p^x q^{n-x} \geq 0$.

$$\sum_x p_r$$

b. (X) = 1, untuk semua x.

10.3.2 Distribusi Multinomial

Kalau pada distribusi binomial hasil sebuah percobaan hanya dikatagorikan 2 macam, yaitu sukses dan gagal, maka dalam distribusi multinomial, sebuah percobaan akan menghasilkan beberapa kejadian (lebih dari 2) yang saling meniadakan atau saling lepas. Misalkan ada sebanyak k kejadian dalam sebuah percobaan, misalnya kejadian B1, B2, …, Bk. Jika percobaan diulang sebanyak n kali dan peluang terjadinya setiap kejadian B konstan dari setiap percobaan dengan P(Bi) = Pi untuk i = 1, 2, 3, …, k, dan X1, X2, X3, …Xk menyatakan jumlah terjadinya kejadian Bi (i = 1, 2, …, k dalam n percobaan). Fungsi distribusi multinomial, adalah:

$$p(x_1, x_2, x_3, \ldots\ldots, x_k) = \left[\frac{n!}{x_1!x_2!x_3!\ldots\ldots x_k!}\right] p_1^{x_1} p_2^{x_2} p_3^{x_3} \ldots\ldots p_k^{x_k}$$

untuk nilai-nilai

$$X_1 = 0, 1, 2, \ldots\ldots; X_k = 0, 1, 2, \ldots \sum_{i=1}^{k} x_i \text{ dan } = n$$

Dimana:

$X_1, X_2, \ldots\ldots, X_k$ menyatakan jumlah dari kejadian $B_1, B_2, \ldots\ldots B_k$

n menyatakan jumlah percobaan.

$P_1, p_2, .., p_k$ adalah probabilitas terjadinya kejadian $B_1, B_2, \ldots B_k$

Contoh:

Proses pembuatan pensil dalam sebuah pabrik melibatkan banyak buruh dan proses tersebut terjadi berulang-ulang. Pada suatu pemeriksaan terakhir yang dilakukan telah memperlihatkan bahwa 85% produksinya adalah "baik", 10% ternyata "tidak baik tetapi masih bias diperbaiki" dan 5% produksinya "rusak dan harus dibuang". Jika sebuah sample acak dengan 20 unit dipilih, berapa peluang jumlah unit "baik" sebanyak 18, unit "tidak baik tetapi bisa diperbaiki" sebanyak 2 dan unit "rusak" tidak ada?

Misalkan:

X1 = banyaknya unit "baik"

X2 = banyaknya unit yang "tidak baik tetapi bias diperbaiki"

X3 = banyaknya unit yang "rusak dan harus dibuang

$X_1 = 18$, $X_2 = 2$, dan $X_3 = 0$ (syarat $x_1 + x_2 + x_3 = n = 20$)

dan $p_1 = 0{,}85$, $p_2 = 0,_1$ dan $p_3 = 0{,}05$ maka:

$p(18, 2, 0) = (0{,}85)^{18} (0{,}1)^2 (0{,}05)^0$

$$\left[\frac{20!}{18!\,2!\,0!}\right]$$

= 190 (0,85)18 (0,01)

= 0,102

Jadi peluangnya sebesar 0,102.

10.3.3 Distribusi Poisson

Distribusi poisson adalah pengembangan dari distribusi binomial yang mampu mengkakulasikan distribusi probabilitas dengan kemungkinan sukses

(p) sangat kecil dan jumlah eksperimen (n) sangat besar. Karena distribusi poisson biasanya melibatkan jumlah n yang besar, dengan p kecil, distribusi ini biasanya digunakan untuk menghitung nilai probabilitas suatu kejadian dalam suatu selang waktu dan daerah tertentu.

Distribusi Poisson memiliki ciri-ciri, yaitu: (1) banyaknya hasil percobaan yang terjadi dalam suatu interval waktu atau suatu daerah tertentu tidak tergantung pada banyaknya hasil percobaan yang terjadi pada interval waktu atau daerah lain yang terpisah; (2) probabilitas terjadinya hasil percobaan selama suatu interval waktu yang singkat atau dalam suatu daerah yang kecil, sebanding dengan panjang interval waktu atau besarnya daerah tersebut dan tidak tergantung pada banyaknya hasil percobaan yang terjadi di luar interval waktu atau daerah tersebut; dan (3) probabilitas lebih dari satu hasil percobaan yang terjadi dalam interval waktu yang singkat atau dalam daerah yang kecil dapat diabaikan. Peubah acak diskrit X mengikuti distribusi poisson, jika fungsi peluangnya berbentuk:

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

Dimana:

Nilai x = 0, 2, 3, 4, ... , N.

e = sebuah bilangan konstan yang besarnya e = 2,7183 dan λ = konstan.

Ternyata distribusi poisson merupakan parameter: μ = λ dan σ = ϒλ. Distribusi poisson digunakan untuk menentukan peluang suatu peristiwa dalam kesempatan tertentu yang terjadinya sangat jarang. Misalnya pada setiap semester jarang ada mahasiswa yang lupa membayar UKT (Uang Kuliah Tunggal).Distribusi poisson juga dapat pula dianggap sebagai pendekatan kepada distirbusi binomial. Jika dalam distribusi binomial N cukup besar, sedangkan π = peluang terjadinya A, sangat dekat dengan nol sedemikian hingga λ = Np harganya tetap, maka distribusi binomial sangat baik didekati oleh distribusi poisson untuk penggunaannya sering dilakukan pendekatan ini.

Contoh:

Misalnya peluang seseorang akan dapat reaksi buruk setelah minum satu butir obat penenang besarnya 0,0005. Dari 4000 orang yang minum obat penenang tersebut, dapat ditentukan peluang yang mendapatkan reaksi: (a) tidak ada; (b) ada dua orang; dan (c) lebih dari dua orang.

Untuk menentukan peluang tidak ada orang yang mendapat reaksi buruk (dalam hal ini x = 0), dengan menggunakan pendekatan distribusi poisson λ = p = 0,0005 = 2. Jika x = banyaknya orang yang mendapatkan rekasi buruk, maka:

$$P(0) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$
$$= \frac{2{,}7183^{-2}2^{0}}{0!}$$
$$= 0{,}1353$$

Untuk menentukan peluang 2 orang yang mendapatkan reaksi buruk (dalam hal ini x = 2), sehingga:

$$P(0) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$
$$= \frac{2{,}7183^{-2}2^{0}}{2!}$$
$$= 0{,}2706$$

Untuk menentukan peluang lebih dari 2 orang yang mendapatkan reaksi buruk (dalam hal ini x = 3, 4, 5, dan seterusnya). Diketahui P(0) = P(1) = P(2) = P(3) + ... + P(n) = 1 = –P(0) – P(1) – P(2), maka:

$$P(1) = \frac{e^{-1}\lambda^{0}}{1!}$$
$$= \frac{2{,}7183^{-2}2^{0}}{1!}$$
$$= 0{,}2706$$

Sedangkan peluang yang dicari adalah: 1 – (0,1353 + 0,2706 + 0,2706) = 0,3235.

Contoh:
Seorang yang akan menjual mobil mewahnya memasang iklan pada suatu surat kabar yang dapat mencapai 100.000 pembaca. Dengan anggapan nilai probabilitas bahwa seorang yang membaca iklan tersebut berminat akan membeli mobilnya sebesar p = 1/50.000. Jika dari 100.000 pembaca ada dua orang yang berminat membeli mobil tersebut (p = 0,00002) dan X = banyaknya pembaca yang berminat pada mobil tersebut, berapakah P(X = 0), P(X = 1), P(X = 2), P(X = 3), P(X = 4)?

Persoalan ini sebetulnya dapat dipecahkan dengan menggunakan fungsi Binomial, karena persoalannya hanya mencari probabilitas x "sukses" dari n = 100.000 eksperimen, dimana probabilitas sukses p = 1/50.000. Akan tetapi karena n terlalu besar dan p terlalu kecil, fungsi poisson dapat digunakan sebagai suatu pendekatan yang lebih sederhana. Apabila l = rata-rata distribusi = E(X) = np = 100.000/50.000 = 2, (secara rata-rata dapat diharapkan dua orang pembaca yang menanyakan keadaan mobil), maka setelah dilakukan perhitungan, akan memperoleh:

$$P_r(X=0) = \frac{2^0 e^{-2}}{0!} = 0{,}1353 \qquad P_r(X=5) = \frac{2^5 e^{-2}}{5!} = 0{,}0361$$

$$P_r(X=1) = \frac{2^1 e^{-2}}{1!} = 0{,}2707 \qquad P_r(X=6) = \frac{2^6 e^{-2}}{6!} = 0{,}0120$$

$$P_r(X=2) = \frac{2^2 e^{-2}}{2!} = 0{,}2707 \qquad P_r(X=7) = \frac{2^7 e^{-2}}{7!} = 0{,}0034$$

$$P_r(X=3) = \frac{2^3 e^{-2}}{3!} = 0{,}1804 \qquad P_r(X=8) = \frac{2^8 e^{-2}}{8!} = 0{,}0009$$

$$P_r(X=4) = \frac{2^4 e^{-2}}{4!} = 0{,}0902 \qquad P_r(X=9) = \frac{2^9 e^{-2}}{9!} = 0{,}0002$$

Perhitungan ini dapat juga dilihat pada Tabel Poisson, dimana x = 0, 1, 2, …, 9. Misalnya ingin melihat distribusi probabilitas bahwa 5 orang pembaca berminat pada mobil tersebut p(5) dengan λ atau rata-rata distribusi = 2, perhatikan potongan Tabel Poisson berikut.

x	1,0	2,0	3,0	4,0	5,0	6,0
0	0,3679	0,1353	0,0498	0,0183	0,0067	0,0025
1	0,3679	0,2707	0,1494	0,0733	0,0337	0,0149
2	0,1839	0,2707	0,2240	0,1465	0,0842	0,0446
3	0,0613	0,1804	0,2240	0,1954	0,1404	0,0892
4	0,0153	0,0902	0,1680	0,1954	0,1755	0,1339
5	0,0031	**0,0361**	0,1008	0,1563	0,1755	0,1606
6	0,0005	0,0120	0,0504	0,1042	0,1462	0,1606

Perhatikan kolom 2, dengan $\lambda = 2,0$, telusuri ke bawah sampai ke baris x = 5. Berdasarkan Tabel Poisson akan ditemukan angka 0,0361. Artinya probabilitas 5 orang berminat dari 100.000 pembaca adalah 0,0361; probabilitas 6 orang berminat adalah 0,0120; dan seterusnya. Distribusi Poisson juga dapat digunakan untuk menghitung probabilitas dari x "sukses" dalam n eksperimen, yang terjadi dalam satuan luas tertentu, satuan isi tertentu, interval waktu tertentu, atau satuan panjang tertentu.

10.3.4 Distribusi Hipergeometrik

Distribusi hipergeometrik sangat erat kaitannya dengan distribusi binomial. Perbedaan antara distribusi hipergeometrik dengan binomial adalah bahwa pada distribusi hipergeometrik, percobaan tidak bersifat independen. Artinya antara percobaan yang satu dengan yang lainnya saling berkait. Selain itu probabilitas "sukses" berubah (tidak sama) dari percobaan yang satu ke percobaan lainnya. Untuk mencari probabilitas x sukses dalam ukuran sample n, peneliti harus memperoleh x sukses dari r sukses dalam populasi, dan n – x gagal dari N – r gagal. Sehingga fungsi probabilitas hipergeometrik dapat dituliskan:

$$p(x) = \frac{{}^{r}C_{x}\ {}^{N-r}C_{n-x}}{{}^{N}C_{n}}, 0 \leq x \leq r$$

Keterangan:

p(x) = probabilitas x sukses (jumlah sukses sebanyak x) dalam n percobaan
n = jumlah percobaan
N = Jumlah elemen dalam populasi
r = jumlah elemen dalam populasi berlabel "sukses"
x = Jumlah elemen berlabel "sukses" diantara n elemen percobaan

Terdapat dua persyaratan yang harus dipenuhi oleh sebuah distribusi hipergeometrik, yaitu: (1) percobaan diambil dari suatu populasi yang terbatas, dan percobaan dilakukan tanpa pengembalian; dan (2) ukuran sampel n harus lebih besar dari 5% dari populasi N. Dari rumus di atas, perhatikan bahwa:

$$^{r}C_{x}=\frac{r!}{x!(r-x)!}$$

$$^{N-r}C_{n-x}=\frac{(N-r)!}{(n-x)(N-r-n+x)!}$$

$$^{N}C_{n}=\frac{N!}{n!(N-n)!}$$

Contoh:

Sebuah anggota komite terdiri dari 5 orang, dimana 3 adalah wanita dan 2 laki-laki. Misalkan 2 orang dari 5 orang anggota komite tersebut dipilih untuk mewakili delegasi dalam sebuah konvensi. Berapa probabilitas bahwa dari pemilihan secara acak didapat 2 orang wanita? Berapa probabilitas dari 2 orang yang terpilih adalah 1 laki-laki dan 1 wanita?

Jawaban

Penyelesaian soal ini dapat menggunakan distribusi hipergeometrik, dengan n = 2; N = 5; r = 3; dan x = 2, x = jumlah wanita terpilih.

Soal 1

$$\mathrm{p}(2)=\frac{^{3}C_{2}\ ^{2}C_{0}}{^{5}C_{2}}=\frac{\left(\frac{3!}{2!1!}\right)\left(\frac{2!}{2!0!}\right)}{\left(\frac{5!}{2!3!}\right)}=\frac{3}{10}=0{,}3$$

Jadi probabilitas 2 orang wanita terpilih adalah 0,3.

Soal 2

$$\mathrm{p}(1)=\frac{^{3}C_{1}\ ^{2}C_{1}}{^{5}C_{2}}=\frac{\left(\frac{3!}{1!2!}\right)\left(\frac{2!}{1!1!}\right)}{\left(\frac{5!}{2!3!}\right)}=\frac{3.2}{10}=\frac{6}{10}=0{,}6$$

Jadi probabilitas terpilih 1 orang wanita dan 1 laki-laki = 0,6.

10.4 DISTRIBUSI PELUANG KONTINU

Sama halnya dengan distribusi peluang diskrti, kita akan menghitung nilai peluang dari suatu pemboleh ubah acak X dimana sekarang nilai dari peubah acaknya adalah kontinu. Distribusi peluang untuk peubah acak kontinu yang akan dibahas adalah Distribusi Normal, Distribusi Studen (t), Distrubsi *chi-square*, dan Distribusi Fisher (F). Hampir setiap rumusan statistik dikembangkan dengan pengasumsian distribusi Normal. Selain itu, distribusi normal sangat banyak digunakan oleh peneliti dan praktisi untuk menyelesaikan permasalahan secara statistik, terutama yang berkaitan denan bentuk distribusi nilai data.

10.4.1 Distribusi Normal

Distribusi normal sering juga disebut Distribusi Gauss. Distribusi normal merupakan salah satu distribusi yang paling banyak digunakan dalam praktik penelitian di lapangan. Beberapa ciri-ciri distribusi normal adalah: (1) grafiknya selalu berada di atas sumbu mendatar X; (2) memiliki bentuk simetris terhadap $X = \mu$; (3) ujung-ujung grafiknya berasimtut terhadap sumbu mendatar X, dimulai dari $X - 3\sigma$ ke kiri dan $X + 3\sigma$ ke kanan; dan (4) luas daerah grafik selalu sama dengan satu satuan luas. Bila harga σ semakin kecil maka kurvanya semakin menjulang tinggi (leptokurtik) dan sebaliknya bila harga σ semakin besar maka kurvanya semakin menjulang bawah (platikurtik).

Distribusi normal merupakan distribusi teoritik dari peubah random kontinum. Distibusi normal adalah distribusi kontinu yang sangat penting dalam statistik dan banyak dipakai memecahkan persoalan. Misalnya hitung $P(X < 1{,}25)$ pada taraf signifikansi 0,05? Cara penyelesaian adalah lihat pada Tabel Distribusi Normal (Lampiran 4), carilah angka 1,2 pada kolom paling kiri. Selanjutnya, carilah angka 0,05 pada baris paling atas. Sel para pertemuan kolom dan baris tersebut adalah 0,8944. Dengan demikian $P(X < 1{,}25)$ adalah 0,8944.

10.4.2 Distribusi t (Student Distribution)

Bentuk grafik distibusi t mirip dengan distribusi normal baku, hanya saja bentuknya lebik leptokurtik, distribusi t yang memiliki derajat kebebasan (db) = $n - 1$. Bila jumlah populasi $N \geq 30$, distribusi t dapat didekati dengan distribusi normal. Untuk menghitung setiap bagian di bawah kurva telah disediakan tabel distribusi t sebagaimana pada Lampiran 9. Distribusi t selain digunakan untuk menguji suatu hipotesis, juga untuk membuat pendugaan interval. Lazimnya distribusi t digunakan untuk menguji hipotesis mengenai nilai parameter, paling banyak dari 2 populasi (lebih dari 2, harus digunakan F), dan dari sampel yang kecil, misalnya n < 100, bahkan seringkali n ≤ 30. Untuk n yang cukup besar (n ≥ 100, atau mungkin cukup n > 30) dapat digunakan distribusi normal, maksudnya tabel normal dapat digunakan sebagai pengganti tabel t. Kalau Z = N(0,1) = peubah normal dengan rata-rata 0 dan simpangan baku 1, dam = kai-kuadrat dengan derajat kebebasan υ, maka peubah t dapat diperoleh dengan rumus:

$$t = \frac{Z}{\sqrt{\frac{\chi_{\upsilon}^{2}}{\upsilon}}}$$

Artinya, fungsi mempunyai distribusi t dengan derajat kebebasan sebesar υ. Peubah t dapat mengambil nilai negatif maupun positif, oleh karena pada dasarnya peubah t ini berasal dari peubah normal (peubah normal selain mengambil nilai positif, juga negatif). Peubah t juga mempunyai kurva yang simetris terhadap t = 0. Tabel t, seperti tabel distribusi normal, dapat digunakan untuk mencari nilai peubah t apabila nilai probabilitas α sudah diketahui, atau sebaliknya. Untuk menggunakan tabel t harus ditentukan terlebih dahulu besarnya nilai α dan υ. Oleh karena kurva t simetris, maka peneliti boleh hanya mencari nilai t sebelah kanan titik 0. Jika sampel kecil (n < 30), maka S2 akan berfluktuasi cukup besar dari sampel ke sampel, sehingga perlu statistik yang lebih baik. Jika sampel kecil akan tetapi berasal dari distribusi normal, maka rumus statistik t ialah:

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

Contoh:

Seorang peneliti menyatakan rata-rata hasil panen setelah diberi pupuk adalah 500 gram per mm pupuk yang diberikan. Dia kemudian mengambil sampel 25 batch panen, dan memutuskan akan puas dengan klaimnya jika ternyata nilai t dari sampel terletak antara –t0,05 s.d. t0,05. Peneliti mengasumsikan bahwa bobot hasil panen mengikuti distribusi normal. Ternyata sampelnya memiliki rata-rata 518 gram dengan deviasi standar sampel 40. Apakah dia akan puas dengan klaimnya?

Jawaban
Hal ini adalah persoalan distribusi t. Ukuran sampel n = 25, sehingga derajat kebebasan ν = n – 1 = 25 – 1 = 24. Dari tabel diketahui bahwa untuk v = 24, maka t0,05 = 1,711, sedangkan hasil sampelnya adalah:

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

$$= \frac{518-500}{40/\sqrt{25}}$$

$$= 2{,}25$$

Berdasarkan hasil perhitungan dapat diketahui bahwa $t_{tabel} = 1{,}711 < t_{hitung} = 2{,}25$. Sehingga dapat disimpulkan peneliti memiliki cukup bukti bahwa dia dapat merasa puas dengan hasil panennya.

10.4.3 Distribusi *Chi Square*

Distribusi *Chi Square* merupakan distribusi peubah acak kontinu. Grafik distribusi *Chi Square* lazimnya merupakan kurva positif, yaitu miring ke kanan. Distribusi *Chi Square* sangat berguna sebagai kriteria untuk pengujian hipotesis mengenai varians dan juga untuk uji ketepatan penerapan suatu fungsi apabila digunakan untuk data hasil observasi atau data empiris. Dengan demikian, peneliti dapat menentukan apakah distribusi pendugaan berdasarkan sampel hampir sama atau mendekati distribusi teoritis, sehingga peneliti dapat menyimpulkan bahwa populasi dari mana sampel yang dipilih mempunyai distribusi yang dimaksud (misalnya, suatu populasi mempunyai distribusi binomial, poisson, atau normal).

Untuk perhitungan luas setiap bagian kurva telah disediakan tabelnya sebagaimana terlampir pada Lampiran 5. Misalnya mencari luas di bawah kurva

Z untuk $\alpha = 0{,}05$ dan N = 14, dapat diketahui bahwa koefisien *Chi Square* tabel sebesar 23,685.

Contoh:

Misalnya sebuah dadu yang mempunyai 6 mata (mata 1, 2, 3, 4, 5, 6) dilemparkan ke atas sebanyak 300 kali. Dalam jangka panjang, diharapkan untuk melihat masing-masing mata tersebut muncul dengan frekuensi yang sama, yaitu masing-masing muncul 50 kali. Dalam praktiknya, frekuensi mata dadu yang muncul sekitar 50, walaupun dadu itu termasuk "*fair dice*". Dengan menggunakan *Chi Square*, dapat menentukan apakah suatu dadu dapat dikatakan "*fair*" setelah membandingkan frekuensi dari masing-masing mata dadu tersebut. Apabila Z_i = N(0,1) = peubah normal dengan rata-rata 0 dan variens sama dengan 1, atau E(Z) = 0, $\sigma_Z^2 = 1$, maka jumlah $Z_1^2 + Z_2^2 + \dots Z_k^2$ sama dengan X_k^2 dengan derajatkebebasan sebesar k.

$$\chi_k^2 = \sum Z_i^2$$

Kalau suatu himpunan yang terdiri n peubah acak X = {Xi}, dimana $Xi \sim n(\mu, \sigma^2)$ untuk semua i (i = 1, 2, …, n), maka dapat diperoleh peubah Z seperti yang dimaksud di atas, dengan rumus:

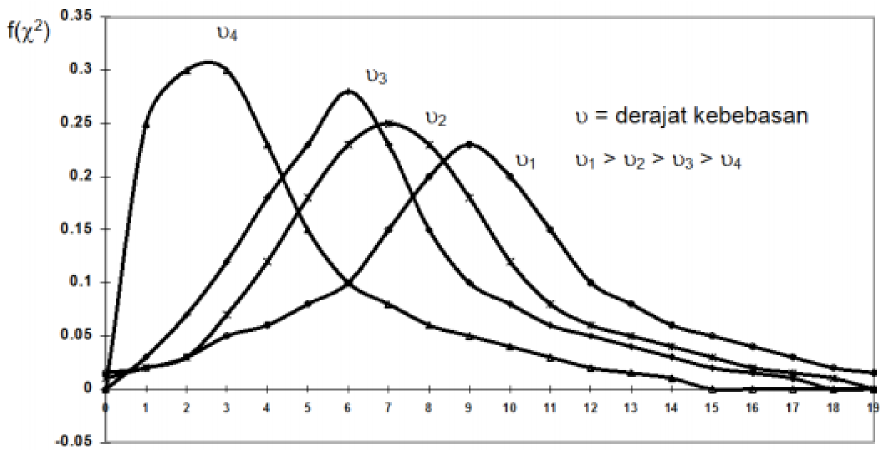
$$Z_i = \frac{X_i - \mu}{\sigma} \sim N\ (90,\ 1);\ i = 1, 2, \dots, n.$$

$$\chi_n^2 = \sum \left(\frac{X_i - \mu}{\sigma} \right)^2$$

χ_n^2 = *chi square* dengan derajat kebebasan sebesar n.

Apakah yang dimaksud dengan derajat kebebasan? Misalnya peneliti diminta untuk menentukan 5 nilai X, yaitu X_1, X_2, X_3, X_4, X_5, dimana syaratnya sudah ditentukan bahwa rata-ratanya. X = 5. Jadi jumlah kelima nilai X tersebut adalah 25. Kalau nilai x_1, x_2, x_3, dan x_4 ditentukan, misalnya $x_1 = 4$, $x_2 = 5$, $x_3 = 6$, dan $x_4 = 7$, maka nilai x_5 tidak bebas lagi untuk menentukannya. Nilai x5 harus membuat jumlah kelima nilai x tersebut menjadi 25. Dengan demikian $X_5 = 25 - (X_1 + X_2 + X_3 + X_4) = 25 - (4 + 5 + 6 + 7) = 3$. Jadi $X_5 = 3$. Sehingga mempunyai 4 kebebasan (satu kali tidak) di antara 5 pilihan, dengan kata lain hanya mempunyai derajat kebebasan sebanyak 4 yaitu (5 – 1).

Kalau harus memilih dari n elemen derajat kebebasannya = n – 1. Di dalam soal hanya memperkirakan satu penduga (X), sehingga derajat kebebasan ada (n – 1). Apabila harus memperkirakan k penduga, maka derajat kebebasan ada (n – k). Kalau k = 2, yaitu a dan b (a penduga dari A dan b penduga dari B), dalam persoalan regresi $\hat{Y} = A + BX + \varepsilon$, maka derajat kebebasannya (n – 2). Bentuk kurva *Chi Square* sangat dipengaruhi oleh besar kecilnya nilai derjat kebebasan. Makin kecil nilai derajat kebebasan, bentuk kurvanya makin menceng kekanan dan makin besar nilai derajat kebebasan ($n \rightarrow \infty$), bentuk kurvanya makin mendekati bentuk funsi normal. *Chi Square* merupakan fungsi kontinu dan nilainya tidak pernah negatif. Nilai rata-ratanya makin meningkat kalau nilai derajat kebebasan juga makin meningkat. Kurva *Chi Square* dengan derajat kebebasan x = (X_x^2).



GAMBAR 10.1 Kurva *Chi Square*

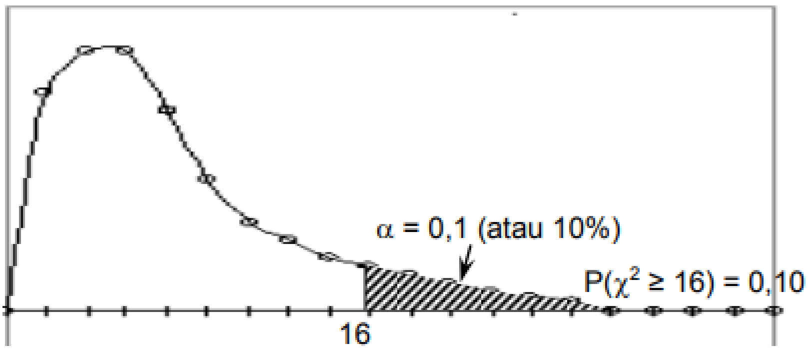
Berdasarkan Gambar 2.7 E (X_υ^2) = $\mu = \upsilon$; rata-rata *Chi Square* dengan derajat kebebasan sebesar υ adalah sama dengan υ; dan Var (X_υ^2) = $\sigma^2 = 2\upsilon$. Untuk keperluan perhitungan nilai X^2, tabel *Chi Square* telah dibuat menurut berbagai nilai derajat kebebasan. Dalam tabel, derajat kebebasan sering diberi simbol υ, r, atau n dan sering disingkat dof atau df. Dalam membaca tabel *Chi Square*, agar diperhatikan simbol (notasi) di bagian atas yang digunakan dalam tabel tersebut. Tabel *Chi Square* memuat nilai X^2, dan bukan nilai probabilitas seperti halnya tabel distribusi normal. Untuk $\upsilon > 100$, distribusi *Chi Square* mendekati distribusi normal, di mana peubah Z sebagai peubah normal baku dapat diperoleh dengan cara:

$$Z = \sqrt{2\chi^2} - \sqrt{2\upsilon - 1}$$

Nilai X^2 dengan nilai derajat kebebasan yang berbeda dan tingkat probabilitas yang berlainan dapat dilihat dalam tabel X^2. Cara membaca Tabel *Chi Square*, misalkan α = probabilitas bahwa *Chi Square* mengambil nilai sama atau lebih besar dari nilai yang terdapat pada tabel *Chi Square* dengan derajat kebebasan sebesar υ. Nilai *Chi Square* dari tabel diberi symbol $X^2_{\alpha,\upsilon}$.

$$P\left(\chi^2 \geq \chi^2_{\alpha,\upsilon}\right) = \alpha$$

Kalau υ= 10, dan α= 10% maka luas daerah dari kurva *Chi Square* berdasarkan tabel *Chi Square* terletak di sebelah kanan dari $X^2_{(0,10)(n)}$ = 16,00.



GAMBAR 10.2 **Luas Daerah Kurva *Chi Square***

Sebagian dari tabel *Chi Square* ditampilkan pada tabel berikut:

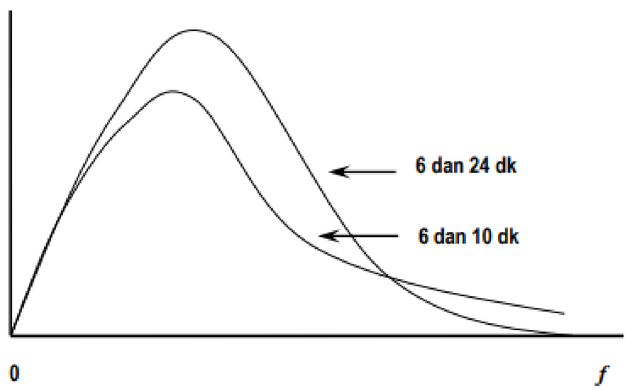
a / b	90% (0,90)	50% (0,50)	10% (0,10)	5% (0,05)
1	0,01	0,45	2,71	3,84
5	1,61	4,35	9,24	11,10
10	4,87	9,34	16,00	18,30
20	12,40	19,30	28,40	31,40
30	20,60	29,30	40,30	43,80
40	29,10	39,30	51,80	55,80

Tabel *Chi Square* memuat nilai hasil perhitungan kumulatif. Misalnya dari tabel di atas, untuk derajat kebebasan sebesar $\upsilon = 10$, luas kurvanya sebesar 90% terletak disebelah kanan titik di mana $X^2 = 4{,}87$.

10.4.4 Distribusi Fisher (F)

Distribusi F juga merupakan distribusi dari peubah acak kontinu. Distribusi F memiliki dua derajat kebebasan, yaitu N1 = derajat kebebasan pembilang dan N2 = derajat kebebasan penyebut. Bentuk grafik distribusi F tidak semitris, tetapi agak positif. Untuk perhitungan Tabel F seperti pada Lampiran 12. Misalnya berapa nilai F untuk $\alpha = 0{,}05$ bila N1 = 15 dan N2 = 25? Untuk keperluan tersebut harus diperiksa derajat kebebasan pembilang 15 di bagian atas dan derajat kebebasan penyebut 25 di bagian bawah. Selanjutnya ditemukan koefisien sebesar 2,09.

Distribusi F adalah prosedur statistika untuk mengkaji (mendeterminasi) apakah rata-rata hitung (mean) dari 3 (tiga) populasi atau lebih, sama atau tidak. Distribusi F digunakan untuk menguji rata-rata atau nilai tengah dari tiga atau lebih populasi secara sekaligus, apakah rata-rata atau nilai tengah tersebut sama atau tidak sama. Kurva distribusi F tidak hanya tergantung pada kedua parameter v1 dan v2, tetapi juga pada urutan keduanya ditulis. Begitu kedua bilangan itu ditentukan, maka kurvanya menjadi tertentu. Gambar 2.9 adalah kurva khas distribusi F.



GAMBAR 10.3 ***Kurva Distribusi F***

Distribusi F akan dipakai untuk memeriksa kesamaan rata-rata dari beberapa group sampel yang diambil secara independen. Ada dua faktor yang akan menentukan apakah perbedaan rata-rata sampel memang nyata atau tidak, yaitu: (1) variasi di dalam sampel (within); dan (2) variasi antar sampel (between).

Bab 11
Sampling

11.1 PENGANTAR

Sebagaimana kita ketahui bahwa statistika terdiri dari dua jenis, yaitu statistika deskriptif dan induktif. Dalam statistika dikenal adanya populasi. Populasi merupakan jumlah nilai keseluruhan yang mungkin, baik dari hasil pengukuran ataupun perhitungan, kualitatif atau kuantitatif, daripada karakteristik sekelompok objek yang lengkap dan jelas. Sedangkan sampel adalah bagian dari seluruh populasi yang diperoleh dengan ketentuan rumusan tertentu. Sehingga, untuk memperoleh kesimpulan yang menyeluruh diperlukan langkah-langkah yang benar untuk pengambilan sampel dalam populasi. Oleh karena itu, bab ini akan menjelaskan cara pengambilan sampel yang sederhana dan akan menguraikan penjelaskan sesuai dengan kaedah statistika.

11.2 DEFINISI SAMPLE

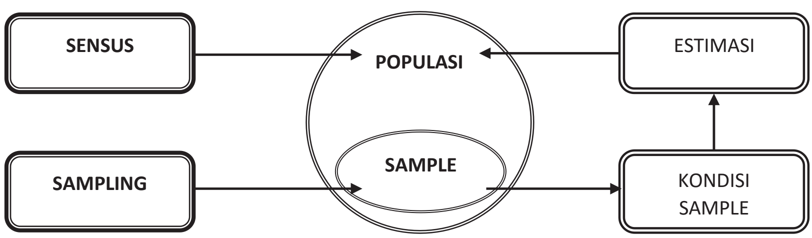
Sampling adalah metode penelitian, yang kesimpulannya terhadap populasi diteliti didasarkan pada hasil pengujian terhadap sampel. **Sampel** adalah merupakan sekelompok pos atau elemen yang diambil dari populasi.

Sampel juga bisa diartikan bagian dari populasi, yang dipilih untuk diteliti, berfungsi sebagai perwakilan dari seluruh anggota populasi. **Populasi**, yang dikenal pula dengan istilah field atau universe, adalah sekumpulan data yang menjelaskan beberapa kejadian yang menjadi perhatian peneliti. Populasi dapat jugadiartikan sebagai kumpulan yang lengkap dari kelompok data yang menjadi objek penelitian. Jadi, dapat disimpulkan bahwa **sampling** adalah penelitian kurang dari 100%, karena dalam sampling, pengujian tidak dilakukan terhadap seluruh anggota populasi. Penelitian 100%, yang menguji seluruh anggota populasi disebut sensus.

Bila metode pengambilan sampel yang dipakai tepat, diharapkan individu-individu sampel yang diobservasi mampu mewakili seluruh anggota populasi dan diperoleh statistik sebagai penduga yang baik Digunakan sampel dalam kegiatan pengumpulan data didasarkan pada berbagai pertimbangan yaitu:

1. Sampel digunakan apabila populasi yang dihadapi terlalu besar, sehingga kecil kemungkinannya untuk diobservasi satu persatu.
2. Pengamatan seluruh anggota populasi dapat bersifat merusak. Misal: bila ingin mengetahui rasa jeruk yang dijual oleh pedagang, tidak mungkin mencicipi semua jeruk dagangannya.
3. Menghemat waktu, biaya dan tenaga.

Hubungan antara sampel, sampling, populasi dan sensus dapat digambarkan sebagai berikut:



Sebagaimana terlihat dalam gambar, pada sampling pengujian dibatasi pada sampel saja. Dari pengujian tersebut diketahui keadaan sampel. Berdasarkan keadaan sampel itu dibuat perkiraan mengenai kondisi populasi yang diuji.

Pekerjaan yang melibatkan populasi tidak digunakan, karena:

1. Mahal dari segi biaya dan waktu yang panjang
2. Populasi akan menjadi rusak atau habis jika disensus

Misal: Dari populasi donat ingin diketahui rasanya, jika semua donat dimakan, dan donat tidak tersisa, tidak ada yang dijual?

Sampel yang baik yaitu Sampel yang representatif

Tabel pembeda antara Statistik Sampel Vs Parameter Populasi

Besaran	Parameter (Populasi)	Statistik (Sample)
Jumlah Observasi	N	n
Rata-rata	μ	X
Varians	σ^2	S^2
Simpang Baku	σ	S
Proporsi	π atau P	P atau $\hat{p}$
Selisih 2 Rata-rata	$\lvert\mu_1 - \mu_2\rvert$: nilai mutlak	$\lvert x_1 - x_2\rvert$: nilai mutlak
Selisih 2 Proporsi	$\lvert\pi_1 - \pi_2\rvert$: nilai mutlak	$\lvert p_1 - p_2\rvert$: nilai mutlak

Sampel yang baik diperoleh dengan memperhatikan hal-hal berikut:

1. Keacakannya (*randomness*)
2. Ukuran
3. Teknik penarikan sampel (*sampling*) yang sesuai dengan kondisi atau sifat populasi
4. Sampel Acak [Random] Dipilih dari populasi di mana setiap anggota populasi memiliki peluang yang sama terpilih menjadi anggota ruang sampel.

11.3 TEKNIK PENGAMBILAN SAMPLE

11.3.1 Probability Sampling

Suatu prosedur pengambilan sampel yang memperhatikan kaidah-kaidah peluang (probability), sehingga bias dan sampling error pengambilan sampel ini dapat ditentukan berdasarkan sampel yang terpilih. Hasil sampling hanya bisa untuk menduga nilai populasinya (parameter).

Macam-macam *Probability Sampling:*

- Sampel Acak Sederhana (Simple Random Sampling), yaitu bila setiap unit dalam populasi diberi peluang sama untuk terpilih. Metode ini merupakan metode yang cukup mudah dan biasa digunakan pada populasi yang memuat karakteristik unit (unit) bersifat relatif homogen. Penarikan Sampel Acak Sederhana (Simple Randomized Sampling), Pengacakan dapat dilakukan dengan: undian, tabel bilangan acak, program komputer.
- Sistimatik Sampling (Systematic Sampling), yaitu suatu metode pengambilan sampel secara random untuk unit sampel yang pertama dan unit-unit sampel selanjutnya dipilih secara sistematik. Penarikan Sampel Sistematik (Systematic Sampling), Tetapkan interval lalu pilih secara acak anggota pertama sampel.

  Contoh:
  Ditetapkan interval = 20
  Secara acak terpilih: Anggota populasi ke-7 sebagai anggota ke-1 dalam sampel maka:
  - ✓ Anggota populasi ke-27 menjadi anggota ke-2 dalam sampel
  - ✓ Anggota populasi ke-47 menjadi anggota ke-3 dalam sampel, dst.

- Sampel Acak Berlapis (Stratified Random Sampling), yaitu metode pemilihan sampel dimana berdasarkan suatu informasi (data) unit-unit di dalam populasi dibuat stratifikasi. Diusahakan nilai-nilai unit di dalam suatu kelompok cukup homogen, sedangkan antar lapisan heterogen. Kemudian dari setiap lapisan yang dibentuk, dipilih sejumlah sampel secara random. Penarikan Sampel Acak Berlapis (Stratified Random Sampling), Populasi terdiri dari beberapa kelas/kelompok. Dari setiap kelas diambil sampel secara acak. Antar Kelas bersifat (cenderung) berbeda nyata (heterogen). Anggota dalam suatu kelas akan (cenderung) sama (homogen).

Contoh:
Dari 1500 penumpang KA (setiap kelas memiliki ukuran yang sama) akan diambil 150 orang sebagai sampel, dilakukan pendataan tentang tingkat kepuasan, maka sampel acak dapat diambil dari:

- ✓ Kelas Eksekutif: 50 orang
- ✓ Kelas Bisnis: 50 orang
- ✓ Kelas Ekonomi: 50 orang

- Sampel Acak Berkelompok (Cluster Sampling), yaitu prosedur sampling di mana unit terkecil dalam populasi merupakan kumpulan dari elemen-elemen. Di dalam cluster biasanya heterogen namun antar cluster homogen. Kemudian kita memilih sebuah sampel yang anggotanya adalah cluster-cluster sehingga bukan lagi sebuah sampel yang anggotanya adalah unit-unit analisis terkecil.

  Penarikan Sampel Gerombol/Kelompok (Cluster Sampling), Populasi juga terdiri dari beberapa kelas/kelompok:

  - ✓ Sampel yang diambil berupa kelompok bukan individu anggota
  - ✓ Antar Kelas bersifat (cenderung) sama (homogen).
  - ✓ Anggota dalam suatu kelas akan (cenderung) berbeda (heterogen).

  Contoh:
  Terdapat 40 kelas untuk tingkat II Jurusan Akuntansi-STEI, setiap kelas terdiri dari 100 orang. Populasi mahasiswa kelas 2, Akuntansi-STEI = 40x100 = 4000. Jika suatu penelitian dilakukan pada populasi tersebut dan sampel yang diperlukan = 600 orang, dilakukan pendataan mengenai lama waktu belajar per hari maka sampel dapat diambil dari 6 kelas. Dari 40 kelas, ambil secara acak 6 kelas.

11.3.1 Non-Probability Sampling

Suatu prosedur pengambilan sampel yang tidak memperhatikan kaidah-kaidah peluang (probability). Biasanya tergantung pada kebijakan dan pengalaman serta subyektifitas dari si peneliti. Bias dan sampling error pengambilan sampel ini tidak dapat ditentukan berdasarkan sampel yang terpilih, sehingga kurang dapat dipertanggungjawabkan untuk analisis secara statistik.

Macam-macam *Non-Probability Sampling:*

- Convinience Sampling, yaitu pengambilan sampel dengan cara ini tidak mewakili secara normal dari target populasi karena unit sampel hanya dipilih berdasarkan conveniently/readily available.
- Judgement/Purposive sampling, yaitu digunakan saat sampel yang diambil berdasarkan pada penilaian yang pasti (expert judgement) mengenai populasi secara keseluruhan (harus mempunyai pengetahuan yang cukup mengenai populasi).
- Quota sampling, yaitu pengambilan sampel dimana jumlah sampel telah ditentukan terlebih dahulu. Pengambil sampel tinggal memilih sampai jumlah tersebut dan biasanya tanpa kerangka sampel. Pengambilan sampel semacam ini sering digunakan dalam public opinion survey.
- Snowball sampling, yaitu digunakan untuk hidden population. Responden diminta memberikan nama dan kontak dari anggota lain dari target populasi. Asumsinya bahwa sesama anggota saling mengenal.

Penarikan Sampel Acak dapat dilakukan dengan 2 cara, yaitu:

1. Penarikan sampel tanpa pemulihan/tanpa pengembalian: setelah didata, anggota sampel tidak dikembalikan ke dalam ruang sampel
2. Penarikan sampel dengan pemulihan: bila setelah didata, anggota sampel dikembalikan ke dalam ruang sampel.

Berdasarkan Ukurannya, maka sampel dibedakan menjadi:

1. Sampel Besar jika ukuran sampel (n) 30
2. Sampel Kecil jika ukuran sampel (n) < 30

11.4 KEMUNGKINAN SAMPEL

Apabila nilai penduga mempunyai kemungkinan cukup besar nilainya akan mendekati nilai populasi, maka tentunya hasil pengambilan sampel yang kita lakukan dapat dikatakan cukup baik, dan kurang baik apabila terjadi sebaliknya. Permasalahannya adalah bagaimana dapat melakukan pengambilan sampel tersebut, sehingga kita bisa memperkirakan tingkat kecermatannya.

Cara yang bisa digunakan adalah dengan menggunakan kaidah-kaidah peluang untuk penarikan unit ke dalam sampel sehingga aspek keacakan dapat terpenuhi. Pada saat menggunakan metode sampel berpeluang, maka setiap menarik unit sebagai anggota sampel, kita tidak mengetahui lebih dahulu unit mana yang akan terpilih. Sebagai gambaran seandainya kita mempunyai 4 unit di dalam populasi, misalnya A, B, C dan D, maka apabila kita gunakan peluang yang sama untuk menarik unit-unit tersebut, masing-masing akan mempunyai peluang yang sama untuk terpilih yaitu ¼.

Bila kita memilih 2 unit sebagai sampel, dan setiap unit dapat terpilih lebih dari sekali, kita dapatkan kemungkinan sampel sebagai berikut:

AA	BA	CA	DA
AB	BB	CB	DB
AC	BC	CC	DC
AD	BD	CD	DD

Jadi seandainya pada penarikan pertama (setelah diundi), kita dapatkan unit C, maka pada penarikan sampel kedua bisa kita dapatkan unit A atau B atau C atau D. Sehingga setiap unit yang terpilih akan mempunyai 4 pasangan yang mungkin. Pasangan-pasangan yang mungkin terpilih dinamakan kemungkinan sampel. Dalam gambaran di atas dinamakan pengambilan sampel dengan ulangan (With Replacement), karena setiap unit bisa terpilih lebih dari sekali. Sehingga banyaknya kemungkinan sampelnya (all possible sample) adalah sebesar $\mathbf{N^n}$.

Jika cara penarikan unit tersebut diubah, yaitu unit yang sudah terpilih tidak boleh dipilih lagi pada pemilihan selanjutnya dan sampel AB dan BA kita anggap sama, maka kemungkinan sampelnya menjadi AB, AC, AD, BC, BD & CD Jadi kita mempunyai 6 kemungkinan contoh. Cara penarikan semacam ini dinamakan penarikan sampel tanpa ulangan (Without Replacement). Sehingga banyaknya kemungkinan sampelnya (all possible sample) adalah sebesar $\mathbf{{}^{N}C_{n}}$.

11.5 KERANGKA SAMPEL

Keseluruhan unit dalam populasi akan membentuk kerangka sampel dan dari sinilah anggota sampel dipilih. Kerangka sampel bisa merupakan daftar dari orang, rumah tangga, perusahaan, catatan dalam sebuah file, kumpulan dokumen, atau berupa sebuah peta dimana telah tergambar unitnya secara jelas. Untuk bisa melakukan penarikan sampel secara acak, kita memerlukan kerangka sampel berupa daftar dari unit berikut keterangan tentang nama, alamat (identifikasi) dan keterangan-keterangan lain yang diperlukan.

Persyaratan yang harus dipenuhi kerangka sampel adalah:

- Memiliki batas yang jelas, artinya setiap unit tidak saling tumpang tindih dengan unit lain.
- Lengkap dan up to date, artinya seluruh unit dalam populasi dalam keadaan terakhir harus didaftar.
- Dapat dikenali, artinya seluruh unit di dalam kerangka sampel dapat dikenal kembali melalui alamat atau petanya.

Bila suatu penarikan sampel dilakukan dalam survei asuransi dengan responden adalah bank, maka kita harus mempunyai kerangka sampel berupa daftar seluruh bank yang ada serta keterangan yang diperlukan dalam wilayah penelitian menurut keadaan terakhir. Bank yang sudah tutup (dilikuidasi) harus dikeluarkan dari kerangka sampel, sedangkan bank yang baru harus dimasukkan ke dalam kerangka sampel lengkap dengan keterangan-keterangan yang diperlukan.

11.6 BESARAN SAMPEL

Ada beberapa hal yang mempengaruhi berapa besaran sampel yang harus diambil dalam sebuah penelitian, sebagai berikut:

1. Heterogenitas dari populasi semakin heterogen. Jika sebuah sampel diambil dari jumlah populasi maka jumlah sampel yang diambil harus semakin besar sehingga karakteristik populasi dapat terwakili.
2. Apabila jumlah peubah yang digunakan semakin banyak maka jumlah sampel yang diambil pun harus semakin besar. Hal ini mengingat adanya persyaratan pengujian hubungan (misalnya dengan *chi-square test of independent* yang tidak memungkinkan adanya sel dengan nilai yang diharapkan < 1 yang dalam perhitungannya dipengaruhi oleh besaran sampel).

3. Teknik penarikan sampel yang digunakan jika menggunakan teknik penarikan sampel acak sederhana maka jumlah sampel tidak terlalu berpengaruh dibandingkan dengan penggunaan teknik penarikan sampel acak berlapis. Semakin banyak lapisan membutuhkan sampel yang besar pula.

Berdasarkan uraikan di atas, adapun rumus yang dapat digunakan untuk menentukan besaran sampel yaitu dengan menggunakan rumus Slovin:

$$n = \frac{N}{1 + Ne^2}$$

n = besaran sampel
N = besaran populasi
e = nilai kritis (batas ketelitian) yang diinginkan (persen kelonggaran ketidaktelitian karena kesalahan penarikan sampel)

Contoh:
Dalam sebuah penelitian, keluarga yang akan diambil sebagai sampel adalah keluarga yang mempunyai kartu sehat. Di daerah kerja Puskesmas Bintulu, terdapat 1.087 keluarga. Dengan menggunakan rumus Slovin di atas, dengan nilai kritis sebesar 10%, maka jumlah sampel yang dibutuhkan dalam penelitian berjumlah 91,57. Oleh karena hasil yang diperoleh dalam bentuk angka diskrit maka angka tersebut harus dibulatkan menggunakan kaedah pembulatan angka. Dalam hal ini 91,57 menjadi 92 keluarga.

Bab 12
Distribusi Sampling

12.1 PENGANTAR

Jika pada bab sebelumnya kita telah membahas apa itu sampling. Pada bab ini akan lebih menekankan pada pendistribusin sampling itu sendiri. Secara defisini distribusi sampling adalah distribusi peluang untuk nilai statistik yang diperoleh dari sampel acak untuk menggambarkan populasi. Sehingga, distribusi sampling adalah distribusi probabilitas dari suatu statistik. Sampling distribution tergantung dari ukuran populasi, ukuran sampel, metode memililih sampel. Distribusi sampling dari $\overline{X}$ dengan dengan ukuran sampel n adalah suatu distribusi yang bila percobaan dilakukan secara berulang (selalu dengan jumlah sampel n) akan menghasilkan banyak nilai sampel dengan rata-rata $\overline{X}$. Distribusi sampling ini menggambarkan variabilitas (perubahan) rata-rata sampel terhadap rata-rata populasi μ.

12.2 DISTRIBUSI PENARIKAN SAMPLE

Nilai-nilai statistik yang dihitung, berdasarkan data sampel biasanya digunakan untuk menduga parameter populasi. Sebagai penduga parameter populasi, tentu saja nilainya tidak selalu persis sama dengan nilai parameter populasi. Jadi ada kemungkinan nilai-nilai statistik lebih besar atau lebih kecil dari nilai parameter populasi.

Ada banyak sekali sampel random yang mungkin dapat diambil dari suatu populasi yang sama. Dengan demikian setiap statistic, misalnya $\bar{x}$ atau S^2 akan berbeda/bervariasi dari satu sampel ke sampel lainnya. Jadi suatu statistic sesungguhnya merupakan variable random yang hanya bergantung kepada sampel yang diamati. Karena statistic merupakan variable random maka tentunya memiliki fungsi distribusi, yang disebut distribusi sampling.

Distribusi sampling terdiri atas:

1) Distribusi Sampling Rata-Rata (*Sample Mean*)
2) Distribusi Sampling Proporsi (*Sample Proportion*)
3) Distribusi Sampling Varian/Ragam (*Sample Variance*)

12.3 DISTRIBUSI SAMPLING RATA-RATA

Misal sampel acak n diambil dari populasi normal dengan rataan μ dan varians a^2. Tiap pengamatan X_i, i = 1, 2, 3, …, n dari sampel acak tersebut akan berdistribusi normal yang sama dengan populasi yang diambil sampelnya. Jadi:

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Berdistribusi normal dengan rataan,

$$\mu_{\bar{X}} = \frac{\mu_1 + \mu_2 + \cdots + \mu_n}{n}$$

Dan, variasi,

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2 + \sigma^2 + \cdots + \sigma^2}{n}$$

Atau,

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

Contoh:

Untuk melihat apakah suatu asset dianggap produktif atau tidak biasanya digunakan nilai return on asset (ROA). Berikut adalah ROA dari beberapa bank relative besar di Indonesia.

BANK	ROA (%)
Bank Bukopin	2
Bank BCA	4
Citi Bank	6
Bang Jabar	4
Bang Tugu	4

a) Hitung berapa nilai rata-rata dan simpangan baku populasi
b) Tentukan Rata-rata dari sampel apabila diambil sampel 2 dari 5 bank yang ada (Tanpa pemulihan)
c) Hitung simpangan baku rata-rata sampel
d) Berapakah nilai peluang rata-rata ROA sampel kurang atau sama dengan 4

Jawaban

a. Rata-rata dan simpangan baku populasi adalah

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{\sum_{i=1}^{5} X_i}{5} = \frac{2+4+6+4+4}{5} = \frac{20}{5} = 4$$

$$\sigma^2 = \frac{\sum_{i-1}^{N} (X_i - \mu)^2}{N} = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2 + (4-4)^2 + (4-4)^2}{5}$$

$$= \frac{8}{5} = 1.6$$

b. Nilai rata-rata dari sampel apabila diambil sampel 2 dari 5 bank yang ada (Tanpa pemulihan), adalah:

1) Jumlah titik sampel rata-rata (Tanpa pemulihan) adalah:

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

12.3.1 Perhitungan Rata-Rata dari Setiap Sampel

		ROA (%)				
		Bukopin	**BCA**	**CitiBank**	**Bank Tugu**	**Bank Jabar**
ROA (%)	Bukopin		2 + 4	2 + 6	2 + 4	2 + 4
	BCA			4 + 6	4 + 4	4 + 4
	CitiBank				6 + 4	6 + 4
	Bang Tugu					4 + 4
	Bank Jabar					

Maka rata-rata sampel-nya adalah:

Bank	Kombinasi ROA (%)	Rata-rata hitung
Bukopin – BCA	2 + 4	(6/2) = 3
Bukopin- Citibank	2 + 6	(8/2) = 4
Bukopin- Bank Jabar	2 + 4	(6/2) = 3
Bukopin – Bank Tugu	2 + 4	(6/2) = 3
BCA – Citibank	4 + 6	(10/2) = 5
BCA – Bank Jabar	4 + 4	(8/2) = 4
BCA – Bank Tugu	4 + 4	(8/2) = 4
Citi Bank – Bank Jabar	6 + 4	(10/2) = 5
Citi Bank – Bank Tugu	6 + 4	(10/2) = 5
Bank Jabar – Bank Tugu	4 + 4	(8/2) = 4

Maka Nilai rata-rata distribusi sampel Rata-rata nya adalah:

$$\bar{X} = \frac{\sum \bar{X}_i}{C_n^N} = \frac{3+4+3+3+5+4+4+5+5+4}{10} = \frac{40}{10} = 4$$

Simpangan baku distribusi sampel rata-rata nya adalah:

X	$(X - \overline{X})$	$\Sigma(X - \overline{X})^2$
3	-1	1
4	0	0
3	-1	1
3	-1	1
5	1	1
4	0	0
4	0	0
5	1	1
5	1	1
4	0	0
Σ		6

$$S_{\bar{x}} = \sqrt{\frac{\sum(\bar{x}_i - \bar{X})^2}{C_n^N - 1}} = \sqrt{\frac{6}{9}} = \sqrt{0.67} = 0.819$$

Distribusi sampel rata-rata nya adalah:

Rata-rata kombinasi	**Frekuensi**	**Probabilitas**
3	3	3/10
4	4	4/10
5	3	3/10
Σ	10	10/10=1

Sehingga Nilai peluang rata-rata ROA kurang atau sama dengan 4 adalah:

$$P(\bar{X} \leq 4) = P(\bar{X} \leq 4) + P(\bar{X} \leq 4) = \frac{3}{10} + \frac{4}{10} = \frac{7}{10} = 0.7$$

1) Distribusi Sampling Rata-rata Sampel Besar

Bila populasi yang disampel tidak diketahui distribusinya, berhingga atau tidak, maka distribusi sampel $\overline{X}$ masih akan berdistribusi hampir normal dengan rataan μ dan varians a^2/n asalkan ukuran sampelnya besar. Ini merupakan akibat dari Teorema Limit Pusat.

DALIL - 1

JIKA …….
Sampel: berukuran = n ≥ 30   rata-rata = $\bar{x}$ $\Big\}$ diambil DENGAN PEMULIHAN dari
$\Big\{$ Populasi berukuran = N   Terdistribusi NORMAL   Rata-rata = μ ; simpangan baku = σ
MAKA………
Distribusi Rata-rata akan mendekati distribusi Normal dengan :
$\mu_{\bar{x}} = \mu$ dan $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ dan nilai $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

DALIL - 2

JIKA …….
Sampel: berukuran = n ≥ 30   rata-rata = $\bar{x}$ $\Big\}$ diambil TANPA PEMULIHAN dari
$\Big\{$ Populasi berukuran = N   Terdistribusi NORMAL   Rata-rata = μ ; simpangan baku = σ
MAKA……….
Distribusi Rata-rata akan mendekati distribusi Normal dengan :
$\mu_{\bar{x}} = \mu$ dan $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}$ dan nilai $z = \dfrac{\bar{x} - \mu}{(\sigma/\sqrt{n})\sqrt{\dfrac{N-n}{N-1}}}$

- $\sqrt{\dfrac{N-n}{N-1}}$ disebut sebagai FAKTOR KOREKSI populasi terhingga.
- Faktor Koreksi (FK) akan menjadi penting jika sampel berukuran n diambil dari populasi berukuran N yang terhingga/terbatas besarnya
- Jika sampel berukuran n diambil dari populasi berukuran N yang sangat besar maka FK akan mendekati $1 \rightarrow \sqrt{\dfrac{N-n}{N-1}} \approx 1$, hal ini mengantar kita pada dalil ke-3 yaitu: DALIL LIMIT PUSAT = DALIL BATAS TENGAH *(CENTRAL LIMIT THEOREM)*

DALIL - 3: DALIL LIMIT PUSAT

JIKA….

$$\left.\begin{array}{l}\text{Sampel:}\\ \text{berukuran} = n\\ \text{rata-rata} = \bar{x}\end{array}\right\}\text{ diambil dari } \left\{\begin{array}{l}\text{Populasi berukuran} = N \text{ yang BESAR}\\ \text{distribusi : SEMBARANG}\\ \text{Rata-rata} = \mu\text{ ; simpangan baku} = \sigma\end{array}\right.$$

MAKA…….

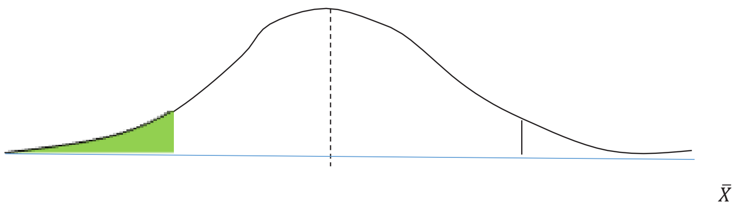
Distribusi Rata-rata akan mendekati distribusi Normal dengan :

$$\mu_{\bar{x}} = \mu \qquad \text{dan} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad \text{dan nilai } z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

- Dalil Limit Pusat berlaku untuk:
  1. penarikan sampel dari populasi yang sangat besar,
  2. distribusi populasi tidak dipersoalkan
- Beberapa buku mencatat hal berikut: Populasi dianggap BESAR jika ukuran sampel KURANG

  DARI 5 % ukuran populasi atau, $\dfrac{n}{N} < 5\%$

Contoh: Sebuah perusahaan lampu pijar menyatakan bahwa lampu pijar yang diproduksi mempunyai usia pemakain tersdiribusi normal dengan rata-rata 800 jam dan deviasi standar 40 jam. Tentukan probabilitas bahwa suau sampel acak dari 16 buah lampu pijar tersebut memiliki rata-rata usia kurang dari 775 jam.



Penyelesaian: distribusi sampling akan mendakti normal dengan dan probabilitas yang diinginkan seperti gambar di atas. Maka:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{775-800}{10} = -2,5$$

$$P(\bar{X} < 775) = P(Z < -2,5) = 0,0062$$

Contoh:

PT AKUA sebuah perusahaan air mineral rata-rata setiap hari memproduksi 100 juta gelas air mineral. Perusahaan ini menyatakan bahwa rata-rata isi segelas AKUA adalah 250 ml dengan standar deviasi = 15 ml. Rata-rata populasi dianggap menyebar normal.

1. Jika setiap hari diambil 100 gelas AKUA sebagai sampel acak DENGAN PEMULIHAN, hitunglah:
   a. Simpangan baku sampel tersebut?
   b. peluang rata-rata sampel akan berisi kurang dari 253 ml?
2. Jika sampel diperkecil menjadi 25 gelas, hitunglah:
   a. Simpangan baku sampel tersebut?
   b. peluang rata-rata sampel akan berisi lebih dari 255 ml?

Jawaban

1. Diselesaikan dengan DALIL 1 karena PEMULIHAN

$$N = 100\ 000\ 000 \qquad \mu_{\bar{x}} = \mu = 250 \sigma = 15 \qquad n = 100$$

$$P(\bar{x} < 253) = P(z < ?)$$

$$Simpangan\ baku = S_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{100}} = \frac{15}{10} = 1.5$$

$$z = \frac{253 - 250}{1.5} = \frac{3}{1.5} = 2.0$$

Jadi $P(\bar{x} > 253) = P(z - 2.0) = 1 - P(z > 2)$

$$= 1 - 0.0228 = 0.9772$$

2. Diselesaikan dengan DALIL 3 karena POPULASI SANGAT BESAR

   $N = 100\ 000\ 000 \qquad \mu_{\bar{x}} = \mu = 250\sigma = 15 \qquad n = 25$

   $P(\bar{x} < 255) = P(z <?)$

$$Simpangan\ baku = S_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{25}} = \frac{15}{5} = 3$$

$$z = \frac{255 - 250}{3.0} = \frac{5}{3.0} = 1.67$$

   Jadi $P(\bar{x} > 255) = P(z > 1.67) = \ldots\ldots\ldots\ldots$(lihat di tabel-z)

12.3.2 Distribusi Sampling Rata-rata Sampel Kecil

DISTRIBUSI t

✓ Distribusi Sampling rata-rata sampel kecil didekati dengan distribusi t Student = distribusi t (W.S. Gosset).

Distribusi-t pada prinsipnya adalah pendekatan distribusi sampel kecil dengan distribusi normal.

Dua hal yang perlu diperhatikan dalam Tabel t adalah

1. derajat bebas (db)
2. nilai α

Derajat bebas (db) = degree of freedom = v = n - 1.

n = ukuran sampel.

Nilai α adalah luas daerah kurva di kanan nilai t atau luas daerah kurva di kiri nilai -t

Nilai $\alpha \rightarrow$ 0.1 (10%) ; 0.05 (5%) ; 0.025(2.5%) ; 0.01 (1%) ; 0.005(0.5%)

Nilai terbatas karena banyak kombinasi db yang harus disusun!

Distribusi t akan gunakan lebih mendalam pada bahasan PENGUJIAN HIPOTESIS

- ✓ Pembacaan Tabel Distribusi-t dapat dilihat pada Tabel dan cara membacanya (terlampir)
- ✓ Perbedaan Tabel z dan Tabel t

  Tabel z→nilai z menentukan nilai α

  Tabel t→nilai α dan db menentukan nilai t

Dalam banyak kasus nilai simpangan baku populasi (σ) tidak diketahui, karenanya nilai σ diduga dari nilai simpangan baku sampel (s)

Dalil 4

JIKA

Sampel:		
Ukuran KECIL n < 30	diambil dari	Populasi berukuran = N
rata-rata = $\bar{x}$ simp. Baku = s		Terdistribusi :Normal
		Rata-rata = μ ; σ = tidak diketahui

MAKA

Distribusi Rata-rata akan mendekati distribusi-t dengan :

$\mu_{\bar{x}} = \bar{X} = \mu$ dan $S_{\bar{X}} = \frac{s}{\sqrt{n}}$ dan nilai $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$

Pada derajat bebas = n-1 dan suatu nilai α

Contoh:

Manajemen PT JURAM menyatakan bahwa 95% rokok produksinya rata-rata mengandung nikotin 1.80 mg, data tersebar normal. Yayasan Konsumen melakukan pengujian nikotin terhadap 9 batang rokok dan diketahui rata-rata sampel = 1.95 mg nikotin dengan standar deviasi = 0.24 mg. Apakah hasil penelitian Yayasan Konsumen mendukung pernyataan Manajemen PT JURAM?

Jawab:

95 % berada dalam selang→berarti 5 % berada di luar selang;

2.5 % di kiri t dan 2.5% di kanan t

α = 2.5 % = 0.025

n = 9→db = n - 1 = 8
t tabel (db, α) = t-tabel(8; 0.025) = 2.306
Jadi 95 % berada dalam selang -2.306 < t < 2.306

Nilai t-hitung = ? $\mu = 1.80 \quad n = 9 \quad \bar{x} = 1.95 \quad s = 0.24$

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = t = \frac{1.95 - 1.80}{0.24/\sqrt{9}} = \frac{0.15}{0.08} = 1.875$$

Nilai t hitung = 1.875 berada dalam selang -2.306 < t < 2.306

Jadi hasil penelitian Yayasan Konsumen masih sesuai dengan pernyataan manajemen PT JURAM.

12.4 DISTRIBUSI SAMPLING PROPORSI (SAMPLE PROPORTION)

Jika sebuah populasi berukuran N di dalamnya terdapat peristiwa A sebanyak Y, maka parameter proporsi peristiwa A sebesar π = Y/N. Dari populasi ini diambil sampel acak berukuran n dan dimisalkan di dalamnya ada peristiwa A sebanyak X, maka proporsi peristiwa A dalam sampel p = X/n.

- Untuk (n/N) >5%, maka digunakan persamaan:

$$\mu_p = \pi$$

$$S_p = \sqrt{\frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right)} \quad \rightarrow \quad Z = \frac{p - \pi}{S_p} \sim N(0,1)$$

- Untuk (n/N) ≤5%, maka digunakan persamaan:

$$\mu_p = \pi$$

$$S_p = \sqrt{\frac{p(1-p)}{n}} \quad \rightarrow \quad Z = \frac{p - \pi}{S_p} \sim N(0,1)$$

Contoh:
Diketahui sebanyak 10% dari ibu-ibu rumah tangga di Bandung memakai detergen A untuk mencuci pakaiannya. Jika dari populasi tersebut diambil sampel berukuran 100:

a. Tentukan rata-rata dan simpangan baku dari populasi ibu-ibu rumah tangga yang memakai detergen A?
b. Bila dari sampel tersebut ternyata terdapat paling sedikit 15 ibu rumah tangga yang memakai detergen A, tentukan probabilitasnya!

Jawaban

1) Rata-rata $= 0.1 \rightarrow p = \mu_p = \pi = 0.1$

$$S_p = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.1(0.9)}{100}} = 0.03$$

2) Proporsi yang memakai detergen A adalah 15/100 = 0,15
Maka: Peluangnya adalah

$$P\left(p > \frac{15}{100}\right) = P\left(Z > \frac{p-\pi}{S_p}\right) = P\left(Z > \frac{0.15-1}{0.3}\right)$$

$$= P(Z > 1.67) = 0.0475$$

12.5 DISTRIBUSI SAMPLING VARIANS (SAMPLE VARIANCES)

- Jika sebuah populasi berukuran N, dari populasi ini diambil sampel acak berukuran n, lalu untuk setiap sampel dihitung simpangan bakunya yaitu S. Dari kumpulan sampel dihitung rata-ratanya yaitu μ_s dan simpangan bakunya σ_s.
- Untuk n ≥ 100, distribusi simpangan baku sangat mendekati distribusi normal dengan Rata-rata $\mu_s = \sigma$ dan simpangan baku $\sigma_s = \frac{\sigma}{\sqrt{2n}}$

$$Z = \frac{S-\mu_s}{\sigma_s} = \frac{S-\sigma}{\left({}^{\sigma}/_{\sqrt{2n}}\right)} = \sim\ N(0.1)$$

Dimana:

S : Simpangan baku sampel

σ : Simpangan baku populasi

σ_s : Simpangan baku dari Simpangan baku

12.6 T-DISTRIBUTION

Teorema: Misalkan Z adalah peubah acak normal standar dan V adalah peubah acak distribusi *chi-squared* dengan derajat kebebasan v. Bila Z dan V adalah independent, maka distribusi peubah acak T,

$$T = \frac{Z}{\sqrt{V/v}}$$

Dengan fungsi kepadatan;

$$h(t) = \frac{\left[\left[\frac{v+1}{2}\right]\right]}{\left[\left[\frac{v}{2}\right]\right]\sqrt{\pi v}}\left(1+\frac{t^2}{v}\right)^{-\frac{v+1}{2}} \quad ; -\infty < t < \infty$$

$$h(t) = \frac{\Gamma[(v+1)/2]}{\Gamma(v/2)\sqrt{\pi v}}\left(1+\frac{t^2}{v}\right)^{-(v+1)/2} \quad ; \quad -\infty < t < \infty$$

Disebut sebagai *t-distribution* dengan derajat kebebasan v.

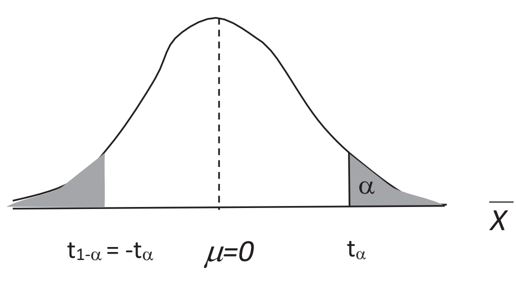
Corollary: Misalkan X_1, X_2,..., X_n adalah peubah acak bebas yang semuanya normal dengan rata-rata μ dan standar deviasi σ. Bila:

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n} \text{ dan } S^2 = \sum_{i=1}^{n} \frac{(Xi-\bar{X})^2}{n-1}$$

Maka variable acak memiliki $T = \frac{\bar{X}-\mu}{s/\sqrt{n}}$ memiliki distribusi-t dengan derajat kebebasan $v = n-1$.

t-distribution sering disebut juga dengan *student t-distribution*.

Distribusi t mirip dengan distribusi normal yakni keduanya simteris dan rata-rata = 0. Karena distribusi-t simetris maka: $t_{1-\alpha} = -t_\alpha$



Contoh: Tentukan nilai-t dengan derajat kebebasan 14 dan luas area a = 0.025 disebelah kiri, atau 0.975 disisi kanan.
Solusi: lihat tabel distribusi t:

$$T_{0.975} = -t_{0.025} = -2.145$$

Contoh:
Tentukan

$$P(-t_{0.025} < T < t_{0.05})$$

Solusi:
Karena luas area disisi kanan = 0.05 dan sebelah kiri 0.025; maka; Luas area T = 1- 0.05 – 0.025 = 0.925, sehingga

$$P(-t_{0.025} < T < t_{0.05}) = 0.925$$

Contoh:
Seorang insinyur kimia menyatakan bahwa rata-rata populasi yang dihasilkan dari suatu proses produksi adalah 500 grams/mm. Untuk membuktikan pernyataan tersebut diambil 25 sampel. Bila nilai t berada antara $-t_{0.05}$ dan $t_{0.05}$ maka pernyataan tersebut benar. Dari hasil percobaan terhadap 25 sampel tersebut menunjukkan bahwa rata-rata $\bar{x}$ =518 gram/mm dan deviasi standar s = 40 gram.

Solusi: dari tabel menunjukkan bahwa untuk derajat kebebasan 24 $t_{0.05}$ = 1.711.

Oleh karena itu, pernyataan tsb benar bila nilai t antara -1.711 dan 1.711. bila μ = 500, maka:

$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

$$T = \frac{518 - 500}{40/\sqrt{25}}$$

$$T = 2.25$$

Probabilitas nilai t sama atau lebih besar dari 2.25 dengan v = 24 adalah mendekati 0.02. Bila m > 500, menghasilkan nilai t yang lebih pantas. Oleh karena itu, hasil produksinya lebih baik dari yang mereka nyatakan.

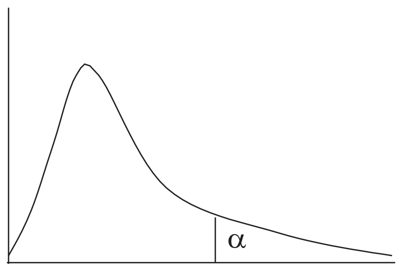
12.7 F-DISTRIBUTION

Teorema: Misalkan *U* dan *V* adalah peubah acak bebas dengan distribusi *chi- squared* dan v_1 & v_2 adalah derajat kebebasan. Maka distribusi peubah acak:

$F = \frac{U/v_1}{V/v_2}$ dengan fu ngsi kepadatan sebagai berikut:

$$h(f) = \begin{cases} \frac{\Gamma[(v_1+v_2)/2](v_1/v_2)^{v_1/2}}{\Gamma(v_1/2)\Gamma(v_2/2)} \cdot \frac{f^{v_1/2-1}}{(1+v_1 f/v_2)^{(v_1+v_2)/2}} & 0 < f < \infty \\ 0 & \text{selainnya} \end{cases}$$

adalah **F-distribution.**



Teorema:

$$f_{1-\alpha}(v_1, v_2) = \frac{1}{f_1(v_2, v_1)}$$

Sehingga nilai *f* dengan 6 dan 10 derajat kebebasan berada pada daerah sisi kanan 0.95 adalah: (lihat tabel distribusi F)

$$f_{0.95}(6{,}10) = \frac{1}{f_{0.05}(10{,}6)} = \frac{1}{4.06} = 0.246$$

Teorema: Bila S_1^2 dan S_2^2 adalah varians dari n_1 dan n_2 sampel acak bebas yang diambil dari populasi normal dengan varians σ_1^2 dan σ_2^2, , maka:

$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\sigma_2^2/S_1^2}{\sigma_1^2/S_2^2}$ adalah Distribusi F

Dengan derajat kebebasan dan $v_1 = n_1 - 1$ dan $v_2 = n_2 - 1$

Bab 13
Teori Pendugaan

13.1 TEORI PENDUGAAN

Suatu Metode statistik dibedakan menjadi dua bagian yaitu Statistik deskriptif (*descriptive* statistics) dan statistic inferensia (*statistical inference*). Statistik deskriptif merupakan suatu metode statistik yang berhubungan dengan pengumpulan data, yang bertujuan untuk memberikan informasi yang akurat. Statistic dekkriptif hanya memberikan data yang yang berupa data sampel atau data populasi tanpa memberikan hasil kesimpulan dari data tersebut. Statistik inferensia merupakan metode statistik yang berisi analisis data sampel atau data populasi yang digunakan untuk menghasilkan suatu peramalan atau prediksi dari suatu data atau populasi. Suatu data atau populasi yang bisa dijadikan contoh adalah pendugaan parameter (*estimate*) dan pengujian hipotesis.

Pendugaan merupakan suatu metode statistik yang digunakan untuk menduga hubungan suatu parameter yang tidak dapat diketahui. Teori Pendugaan banyak digunakan dalam bidang penelitian untuk digunakan sebagai parameter pendugaan suatu penelitian, seperti apakah akan berkaitan dengan suatu sebaran nilai sentral seperti (Rata-rata, median, modus), simpangan atau proporsi. suatu pengetahuan tentang pendugaan sangatlah penting untuk dipelajari. Hasil pendugaan juga harus dipertanggungjawabkan pada suatu penelitian. Dengan menggunakan tingkat kepercayaan dari hasil pendugaan sebagai suatu ukuran untuk menaruh kepercayaan pada ketetapan statistic, maka bisa digunakan untuk menduga parameter populasi.

Teori Pendugaan digolongkan menjadi dua bagian yaitu pendugaan titik dan pendugaan selang. Sedangkan untuk melakukan pendugaan ada berbagai macam cara seperti, menggunakan simpangan kuadrat terkecil, kemungkinan maksimum ataupun sifat pendugaan tak bias linier yang terbaik. Statistic yang digunakan untuk memperoleh nilai dugaan disebut penduga atau fungsi keputusan. Dimisalkan nilai statistik θ atau (*theta*) terdistribusi dengan ciri suatu parameter populasi θ. Parameter θ adalah parameter yang akan diduga dengan nilai dugaan θ yang dapat mengambil bentuk apa saja seperti rata-rata, ragam, simpangan baku atau koefisien regresi.

Untuk menduga sebuah parameter θ perlu dilakukan penarikan contoh yang representatif. Namun sebelum melakukan pendugaan, perlu diketahui terlebih dahulu karakteristik populasi θ seperti bentuk distribusinya. Kita tidak bisa mengharapkan bahwa suatu statistik θ akan menduga θ secara tepat ataupun statistic $\bar{x}$ menduga μ secara eksak, akan tetapi berharap bahwa penduga tidak berada jauh menyimpang dari parameternya.

13.1.1 Cara untuk Menduga

Secara umum penduga μ adalah X yang dirumuskan:

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

Penduga sebuah parameter μ, dimisalkan harganya akan berbeda tergantung pada harga X yang didapatkan dari sampel yang diambil.

13.1.2 Pendugaan Parameter

Pendugaan parameter merupakan cara menduga suatu parameter populasi yang belum diketahui, dengan menggunakan contoh acak dan hitung peluang. Pendugaan pada parameter populasi bisa berupa pendugaan titik atau pendugaan selang (interval). Parameter populasi dan sampel yang digunakan dalam meninjau kelakuan populasi adalag sampel acak. Data dari sampel akan dianalisis, lalu dihitung dan akan memperoleh nilai-nilai statistik yang akan menghasilkan suatu tingkah laku parameter. Harga parameter yang sebenarnya, namun tidak diketahui, akan ditaksir berdasarkan statistik sampel yang diambil dari populasi yang bersangkutan. Parameter populasi yang akan ditaksir dan diuraikan terdiri dari rata-rata, simpangan baku dan proporsi.

Contoh:
Seorang Calon dalam pemilihan Ketua Osis ingin menduga berapa proporsi pemilih yang memilihnya, dengan cara mengambil sampel 200 siswa secara acak untuk ditanyai pendapatnya. Dengan mengetahui proporsi dari pemilih dari calon tersebut akan digunakan sebagai penduga bagi proporsi populasi yang sebenarnya.

13.1.3 Pengujian Hipotesis

Pengujian Hipotesis digunakan untuk memutuskan apakah akan *menerima* atau *menolak* hipotesis tersebut. Dalam pengujian hipotesis ini, sebuah keputusan yang diambil belum tentu benar atau salah, oleh karena itu bisa menimbulkan sebuah resiko. Mengetahui besar kecilnya resiko bisa dinyatakan dalam bentuk probabilitas. Dalam statistic inferensia, pengujian hipotesis sangat dibutuhkan, dalam pengujian ini pembuatan keputusan ataupun dalam pemecah persoalan dalam terselesaikan dengan akurat.

Contoh:
Seorang peneliti arkeolog meneliti sebuah peninggalan sejarah, dengan menggunakan dua jenis alat ukur yang berbeda. Berdasarkan data sampel apakah ada perbedaan ketelitian dari dua jenis alat ukur tersebut, yang diproduksi dari perusahaan yang berbeda.

13.2 PENDUGAAN TITIK (*POINT ESTIMATION*)

Pendugaan titik adalah suatu pendugaan data populasi yang diambil dari data sampel, dengan menyebut hanya satu nilai angka tertentu sebagai estimasi untuk parameter yang tidak diketahui. Pada pendugaan titik ada dua sifat, pertama memiliki nilai harapan penduga yang harus sama dengan parameter yang akan ditaksir. Kedua, memiliki harapan penduga yang harus mempunyai variasi minimum, setiap penduga titik adalah variable random/mempunyai variasi terkecil dari penaksir titik.

Contoh:
Misalkan seorang dosen ingin mengetahui berapakah rata-rata nilai hasil ujian akhir semester (UAS) dari seluruh (N) mahasiswa yang mengikuti matakuliah ekonometrik. Untuk keperluan penelitian, maka diambil sampel (n) sejumlah 15%. Berdasarkan hasil perhitungan yang dilakukan diperoleh nilai rata-rata sebesar 87.

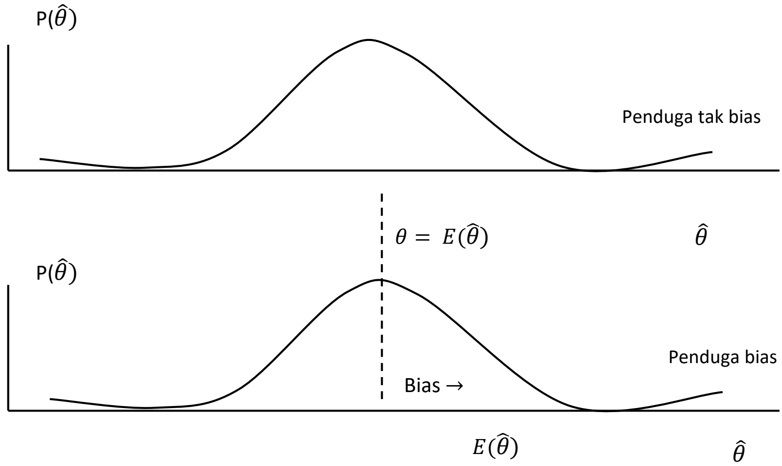
Suatu pendugaan akan menghasilkan bermacam-macam penduga. Diantara penduga-penduga harus dipih mana yang terbaik untuk menduga parameter populasinya. Untuk dapat mengetahui mana yang terbaik dalam memilih penduga, akan disajikan beberapa ciri-ciri penduga yang baik dan penduga yang tidak baik. Penduga yang baik harus memenuhi beberapa syarat, tergantung,dari besar ukuran contohnya. Berikut adalah ciri-ciri penduga yang baik:

a.) Tidak Bias

Penduga parameter θ adalah $\hat{\theta}$ yang merupakan fungsi nilai-nilai penarikan contoh yang diamati yaitu $\hat{\theta} = \hat{\theta}\ (X_1,X_2, \ldots, X_n)$ yang dapat mengambil bentuk apa saja kecuali parameter θ yang tidak diketahui termasuk θ. Distribusi suatu peubah biasanya dikenal memiliki dua ciri dasar yaitu nilai rata-rata populasi dan ragamnya. Nilai rata-rata populasi bagi θ adalah $E(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2$, sedangkan ragamnya $\sigma_{\hat{\theta}}{}^2 = E(\hat{\theta})^2 - [E(\hat{\theta})]^2$. Oleh sebab itu, suatu penduga dikatakan sebagai penduga tak bias bagi parameter jika nilai penduga sama dengan parameter yang diduganya:

$$\mu_{\hat{\theta}} = E(\hat{\theta}) = \theta \qquad (13.2)$$

Nilai penduga $\hat{\theta}$ dikatakan bias, jika $E(\hat{\theta})$ berbeda dari parameter θ seperti yang terlihat pada gambar dibawah ini:



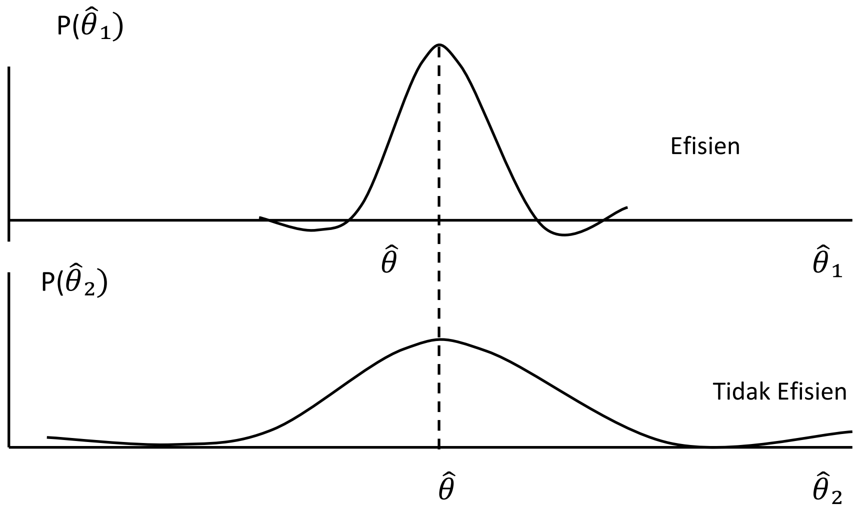
GAMBAR 13.1 ***Perbandingan Penduga tak bias dan Penduga Bias***

Bias $[E(\hat{\theta}) - \theta]$ muncul karena terjadi adanya kesalahan pada waktu melakukan penarikan contoh. Kesalahan ini bisa terdiri dari kesalahan contoh atau kesalahan lainnya. Kesalahan lain bisa dimisalkan kesalahan yang berasal dari manusia, atau karena faktor cuaca atau karena faktor lain yang sulit dihindari. Oleh karena itu, penduga tak bias pun bisa memiliki kesalahan penarikan yang dikarenakan oleh kesalahan lain.

b.) Efisiensi

Suatu penduga dikatakan efisien bagi parameternya, jika penduga memiliki varians yang kecil. Apabila terdapat lebih dari satu penduga, penduga yang efisien adalah penduga yang memiliki varians terkecil. Suatu penduga tak bias, tetapi mempunyai ragam yang besar maka penduga tersebut terdistribusi dengan rata-rata populasi $\hat{\theta}$ namun ragamnya besar sehingga sering mengakibatkan pendugaan yang menyimpang. Maka, bila ragam penduga $\hat{\sigma}_1^2 < \hat{\sigma}_2^2$,maka penduga yang lebih efisien bagi parameter $\hat{\theta}$ adalah ragam $\hat{\sigma}_1^2$.

Dengan memilih penduga yang efisien, pilihlah ragam yang mendekati 0 karena akan memperoleh penduga yang berpusat pada satu titik. Namun jika, untuk mencari penduga yang paling efisien, maka semua penduga tak bias harus dicari ragamnya. Namun jika penduga tak bias efisien jumlahnya sangat banyak, maka akan memakan waktu. Cara yang mudah adalah dengan membandingkan ragam-ragam yang tak bias.



GAMBAR 13.2 ***Pendugaan efisien dan pendugaan tidak efisien (keduanya tidak bias)***

Berdasarkan gambar diatas, penduga $\hat{\theta}^1$ lebih efisien dibandingkan dengan penduga $\hat{\theta}^2$ karena ragamnya lebih kecil.

Efisien relatif terhadap = $\hat{\theta}^2 = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$

Penduga efisien kadang diberikan Batasan, bahwa penduga yang dipilih adalah yang mempunyai ragam paling kecil atau mempunyai kesalahan kuadrat rata-rata minimum. Dengan menggunakan ragam kecil atau mendekati nol, akan diperoleh penduga yang sangat efisien, tetapi tidak berarti tidak mempunyai bias.

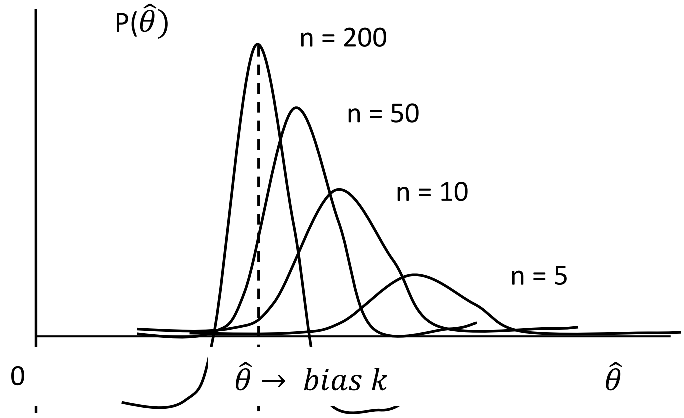
c.) Konsistensi

Penduga parameter $\hat{\theta}$ dikatakan konsisten, apabila nilai dugaan akan menghasilkan hasil yang sama dengan parameter yang diduga dengan bertambahnya ukuran contoh sampai tak terhingga. Secara matematis, penduga $\hat{\theta}$ adalah konsisten jika untuk nilai positif sembarang δ (tidak peduli bagaimana kecilnya) probabilitas mendekati satu:

$$P\left(\left|\hat{\theta} - \theta\right| < \delta\right) =\cong 1$$

$n \rightarrow \infty$

Persamaan diatas memperlihatkan bahwa untuk limit $n \rightarrow \infty$ titik akhir pada saat distribusi penduga $\hat{\theta}$ akan menghampiri titik tertentu yaitu parameter populasinya dengan sekecil mungkin δ (sesaat sebelum titik akhir tercapai).



GAMBAR 13.3 ***Konsistensi penduga untuk n bertambah besar***

$$KKR(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

$$= E[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2$$

$$= E[\hat{\theta} - E(\hat{\theta})]^2 + 2E[\hat{\theta} - E(\hat{\theta})][E(\tilde{\theta}) - \theta] + E[\hat{\theta} - \theta]^2$$

Suku kedua ruas kanan adalah keragaman (covarian) $E[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta] = 0$ sehingga:

$$KKR(\hat{\theta}) = ragam + bias^2 \qquad (13.5)$$

Berdasarkan persamaan (13-3), penduga $\hat{\theta}$ adalah konsisten jika untuk n yang semakin besar atau tidak terhingga KKR($\hat{\theta}$) mendekati nol. Penduga bias yang mendekati nol untuk n tidak terhingga disebut dengan asimtotik tak bias.

13.3 PENDUGAAN SELANG/INTERVAL

Pendugaan selang atau interval adalah suatu pendugaan terhadap parameter populasi yang mempunyai nilai diantara dua nilai. Pendugaan selang atau interval merupakan suatu penduga yang mempunyai pengukuran yang obyektif tentang derajat kepercayaan pada ketelitian pendugaan. Pendugaan selang akan memberikan nilai-nilai statistik dalam suatu interval dan bukan sebagai nilai tunggal pada pendugaan parameter. Dalam pendugaan ini, untuk melihat besarnya nilai parameter yang didugs namun tidak diketahui menggunakan probabilitas.

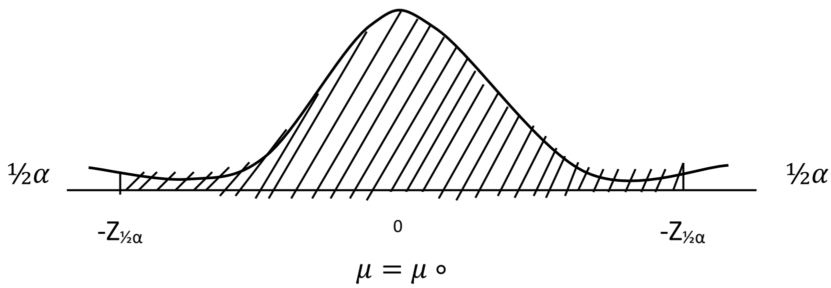
Pendugaan selang/interval memiliki sifat interval kepercayaan atau interval keyakinan dan dibatasi oleh batas keyakinan atas dan keyakinan bawah. Jadi semakin besar penduga interval, maka semakin kecil selang kepercayaannya, sedangkan semakin kecil pendga interval, maka semakin besar selang kepercayaannya. Dipilihlah suatu selang tertentu di sekitar $\hat{\theta}$ yang ditaksir mengandung parameter dan probabilitas tertentu. Untuk membuat pendugaan interval μ harus mendapatkan terlebih dahulu dua nilai statistik $\hat{\theta}^1$ dan $\hat{\theta}^2$ dan menentukan koefisien keyakinan atau tingkat keyakinan yang akan diberikan simbol 1-α.

$$P\left(\hat{\theta}^1 < \theta < \hat{\theta}^2\right) = 1 - \alpha$$

Suatu faktor akan disebut sebagai koefisien tarat nyata (*level of significance,)*. Berdasarkan persamaan diatas, dapat dibentuk suatu probabilitas $\hat{\theta}$ yang terletak pada selang $\hat{\theta}^1 < \theta < \hat{\theta}^2$ yang diharapkan dengan menggunakan derajat kepercayaan (*confidence level*) CL = (1-α) x 100 %, maka akan memberikan nilai parameter dalam suatu penduga selang kepercayaa dan bukan menjadi nilai tunggal. Jika saat terdapat α =0,05, maka akan memperoleh selang kepercayaan sebesar 95% dan jika saat terdapat α = 0,01, maka akan memperoleh selang kepercayaan 99%.

Oleh karena itu, semakin Panjang selang kepercayaan, maka semakin yakin bahwa selang tersebut memuat parameter yang tidak diketahuinya. saat menduga suatu batas-batas selang, dengan derajat kepercayaan yang diambil dari normal baku Z akan memperoleh nilai z negative dan z positif yang mempunyai sifat harga yang mutlak dan memiliki ukuran yang sama besar dengan luas daerah antara keduanya (1-α). Dibawah ini akan disajikan perolehan masing-masing nilai z negative dan z positif dimana luas z ke titik z = 0 dan dinyatakan sebagai -$Z_{½α}$ < Z <$Z_{½α}$ dengan pernyataan probabilitas:

$$P(-Z_{½\alpha} < Z < Z_{½\alpha}) = 1 - \alpha$$



GAMBAR 13.4 ***Probabilitas P(-$Z_{½α}$< Z < $Z_{½α}$) = 1-α***

Contoh:
Misalkan seorang dosen ingin mengetahui berapakah rata-rata nilai hasil ujian akhir semester (UAS) dari seluruh (N) mahasiswa yang mengikuti matakuliah ekonometrik. Untuk keperluan penelitian, maka diambil sampel (n) sejumlah 15%. Berdasarkan hasil perhitungan yang dilakukan diperoleh nilai rata-rata berkisar antara 80 sampai dengan 92.

13.4 PENDUGAAN INTERVAL RATA-RATA

Statistik $\bar{x}$ adalah suatu penduga yang baik untuk parameter populasi μ dengan ragam yang kecil dibandingkan dengan penduga-penduga yang lain. Untuk menurunkan penduga selang bagi nilai rata-rata populasi μ dapat menggunakan teorema limit. Dalam pendugaan rata-rata, ada beberapa yang yang harus diperhatikan seperti Ukuran sampel pada pendugaan rata-rata apakah besar (n>30) atau kecil (n<30). Apakah data terdistribusi normal atau tidak, selain itu juga menentukan standar deviasinya. Jika statistik $\bar{x}$ merupakan rata-rata contoh acak yang berukuran n, yang diambil dari populasi normal dengan simpangan baku $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ yang terletak pada selang $\bar{x}^1$ dan $\bar{x}^2$, maka selang kepercayaan (1 - α)100% populasi μ dinyatakan dalam probabilitas:

$$\mathrm{P}\left(-Z_{½\alpha} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < Z_{½\alpha}\right) = 1 - \alpha$$

Selanjutnya, dikalikan suku-suku ruas kiri dengan $\sigma/\sqrt{n}$, kemudian ditambahkan ruas kiri dalam kurung dengan $-\bar{x}$ sehingga akan diperoleh:

$$P\left(-\bar{x} - Z_{½\alpha}\,\frac{\sigma}{\sqrt{n}} < -\mu < -\bar{x} + Z_{½\alpha}\,\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Selanjutnya dikalikan ruas kiri dalam kurung -1 sehingga tanda pertidaksamaan berubah. Kemudian diatur kembali letak kedua ruas ujung sehingga dapat dipertukarkan:

$$P\left(\bar{x} - Z_{½\alpha}\,\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{½\alpha}\,\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Contoh:

Terdapat 49 mahasiswa korea semester 2 yang mempelajari Bahasa Indonesia, diketahui bahwa rata-rata IPK = 2,9 dengan simpangan baku = 0,3. Berapakah selang kepercayaan 95% untuk rata-rata IPK seluruh mahasiswa dari korea semester 2?

Selang kepercayaan 95% = α = 5% menjadi α /2 = 2.5%

$$Z_{2.5\%} = Z_{0.025} = 1.96$$

$$P\left(\bar{x} - Z_{½\alpha}\,\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{½\alpha}\,\frac{\sigma}{\sqrt{n}}\right)$$

$$\left(2.9 - (1.96)\,\frac{0.3}{\sqrt{49}} < \mu < 2.9 + (1.96)\,\frac{0.3}{\sqrt{49}}\right)$$

$$(2.9 - 0.084 < \mu < 2.9 + 0.084)$$

$$2.816 < \mu < 2.984 \text{ atau } 2.7 < \mu < 3.0 \text{ (dibulatkan)}$$

13.4.1 Selang Kepercayaan μ Jika σ Diketahui

Pendugaan parameter populasi dengan μ statistic $\bar{x}$, pada umumnya ragam populasinya tidak diketahui. Suatu besarnya selang kepercayaan tetap dapat ditentukan setelah dinyatakan derajat atau koefisien kepercayaannya. rumus dibawah ini, berlaku untuk sampel ukuran besar (n >30) dari populasi yang tak terbatas atau dari populasi terbatas tetapi $\frac{n}{N}$:

$$\bar{x} - Z_{½\alpha}\,\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{½\alpha}\,\frac{\sigma}{\sqrt{n}} \qquad (13.11)$$

Z½ adalah nilai Z yang luas daerah yang terletak disebelah kanan dan sebelah kiri dibawah kurva normal baku yaitu $½\alpha$. Langkah selanjutnya, mengganti simpangan baku populasi σ dengan titik taksiran s simpangan baku contohnya. Maka selang kepercayaan dapat diturunkan menjadi:

$$\bar{x} - Z_{½\alpha}\,\frac{s}{\sqrt{n}} < \mu < \bar{x} + Z_{½\alpha}\,\frac{s}{\sqrt{n}}$$

Namun jika populasi N terlalu besar bahkan sampai tak terhingga, maka sebaiknya digunakan faktor koreksi [(N-n)/(N-1)}\]½.

$$\bar{x} - Z_{½\alpha}\,\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} < \mu < \bar{x} + Z_{½\alpha}\,\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Rumus diatas (13.13) berlaku untuk sampel sangat besar (n>30) dari populasi terbatas dengan $\frac{n}{N} > 0.05$.

13.4.2 Selang Kepercayaan Jika Tidak Diketahui

Dengan cara yang sama seperti pendugaan selang kepercayaan sebelumnya, luas dibawah lengkungan kurva distribusi t bergantung pada besar derajat bebas df. Jika $\bar{x}$ dan s yang merupakan rataan dan simpangan baku sampel acak dari populasi normal dengan variansi σ^2 tidak diketahui, selang kepercayaan (1- α)100% untuk μ adalah:

$$\bar{x} - t_{½\alpha}\,\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{½\alpha}\,\frac{s}{\sqrt{n}}$$

bentuk persamaan diatas merupakan pendugaan parameter populasi μ dengan ukuran sampel kecil (n<30), derajat bebas df dan selang kepercayaan (1-

α)100%. Oleh karena itu, $t_{½}$ merupakan nilai t dengan derajat bebas yang terletak disebelah kiri dan di kanannya terdapat daerah seluas ½α.

Contoh:
Suatu perusahaan swasta diteliti untuk mengetahui berapa modal perusahaan swasta di Jakarta. Sampel yang diambil adalah 450 perusahaan swasta dijakarta. Seorang pejabat perbankan menganalisa bahwa dari 90 perusahaan swasta dijakarta, modal perusahaan swasta mencapai Rp. 850 juta. Standar deviasi modal tersebut mencapai 50 juta. Dengan tingkat keyakinan sebesar 99%, berapakah taksiran rata-rata modal perusahaan swasta dijakarta?

Jawab:
Diketahui:

$N = 450$

$n = 90$

$\bar{x} = 850$

$\sigma = 50$

$\frac{n}{N} = \frac{90}{450} = 0{,}2$

$\alpha = 100\% - 99\% = 1\% = 0{,}01 \rightarrow 1\text{-}\alpha = 0{,}99 \rightarrow \frac{1-\alpha}{2} = 0{,}495$

$Z½\alpha = Z_{0,495} \rightarrow Z = 2{,}58$

$$\bar{x} - Z_{½\alpha} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} < \mu < \bar{x} + Z_{½\alpha} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

$$850 - 2{,}58 \frac{50}{\sqrt{90}} \sqrt{\frac{450-90}{450-1}} < \mu < 850 + 2{,}58 \frac{50}{\sqrt{90}} \sqrt{\frac{450-90}{450-1}}$$

$$850 - 12{,}1757727971 < \mu < 850 + 12{,}1757727971$$

$$837{,}8242272029 < \mu < 850 + 862{,}1757727971$$

Maka dari itu, dengan tingkat signifikasi 1%, rata-rata modal perusahaan swasta dijakarta berkisar antara Rp. 837,8242272029 juta atau Rp 837.824.227.202,9 juta dan Rp. 862,1757727971 juta atau Rp 862.175.772.797,1.

13.5 PENDUGAAN INTERVAL BEDA DUA RATA-RATA

Jika mempunyai dua parameter populasi μ_1 dan μ_2 dengan ragam masing-masing σ_1 dan σ_2 dalam suatu pengamatan, maka dapat menduga beda dua rata-rata dengan menghitung selisih statistic $\overline{x_1}$ dan $\bar{x}_2$. Pendugaan interval beda dua rata-rata ditentukan sebagai berikut:

13.5.1 Untuk Sampel Besar dan σ_1 dan σ_2 Diketahui

Dimisalkan ada dua populasi yang berukuran besar, yang pertama mempunyai rata-rata μ_1 dan ${\sigma_1}^2$ dan populasi yang kedua mempunyai rata-rata μ_2 dan ragam ${\sigma_2}^2$. Dari setiap populasi, diambillah contoh acak untuk mendapatkan nilai dugaan bagi selisih μ_1 dan μ_2. Jika ukuran contoh bebas n_1 dan n_2 diambil dari populasi, maka distribusi beda rata-rata statistik $\overline{x_1}$ dan $\overline{x_2}$ akan menyebar menghampiri distribusi normal dengan rata-rata dugaan $\mu\bar{x}_1$ - $\bar{x}_2$ = μ_1 dan μ_2 dan simpangan baku seperti persamaan dibawah ini:

$$\sigma(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{{\sigma_2}^2}{\sigma_2} + \frac{{\sigma_2}^2}{\sigma_2}}$$

Selang kepercayaan (1-α)100% bagi beda rata-rata populasi μ_1 dan μ_2 adalah:

$$(x_1 - x_2) - z½\alpha \sqrt{\frac{{\sigma_1}^2}{n^1} + \frac{{\sigma^1}^2}{n^2}} < \mu^1 - \mu^2 < (x_1 - x_2)$$

$$+ z½\alpha \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} \qquad (13.15)$$

Contoh:

50 orang Karyawan perusahaan swasta asing memiliki upah mingguan rata-rata Rp 350.000,00 dengan simpangan baku Rp 25.000,00. Sedangkan 50

orang karyawan perusahaan nasional memiliki upah mingguan rata-rata Rp 150.000,00 dengan simpangan baku Rp 15.000,00. Dengan menggunakan interval keyakinan 99%, buatlah pendugaan beda rata-rata upah karyawan perusahaan swasta asing dengan perusahaan nasional!

Jawab:

$n_s = 50$ $n_D = 50$

$\bar{x}_1 = 350.000$ $\bar{x}_2 = 150.000$

$\sigma_1 = 25.000$ $\sigma_2 = 15.000$

$1\text{-}\alpha = 99\%$

$\alpha = 1\% = 0{,}005$

$Z_{½\alpha} = Z_{0,005} = 2{,}58$

$$\sigma(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{{\sigma_2}^2}{\sigma_2} + \frac{{\sigma_2}^2}{\sigma_2}}$$

$$= \sqrt{\frac{25.000^2}{50} + \frac{15.000^2}{50}}$$

$$= \sqrt{\frac{625.000.000}{50} + \frac{225.000.000}{50}}$$

$$= \sqrt{12.500.000 + 4.500.000}$$

$$= \sqrt{17.000.000} = 4.123{,}1$$

$\bar{x}_1 - \bar{x}_2 = 350.000 – 150.000$

$= 200.000$

$$(x_1 - x_2) - z½\alpha \sqrt{\frac{\sigma_1{}^2}{n^1} + \frac{\sigma 1^2}{n^2}} < \mu^1 - \mu^2 < (x_1 - x_2) + z½\alpha \sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}}$$

$$200.000 - (2{,}58)\,(4.123{,}1) < \mu_1 - \mu_2 < 200.000 + (2{,}58)(4.123{,}1)$$

189.362,402 < $\mu_1 - \mu_2$ < 210.637,598

Jadi perbedaan rata-rata upah karyawan perusahaan swasta asing dan perusahaan nasional berkisar 189.362,402 sampai 210.637,598

13.5.2 Untuk Sampel Kecil dan $\sigma_1{}^2$ dan $\sigma_2{}^2$ yang tidak diketahui

Jika Ragam $\sigma_1{}^2$ dan $\sigma_2{}^2$ boleh dianggap sama (ukuran n<30), maka pendugaan tetap dapat dilakukan asal distribusi populasi berbentuk genta, dan selang kepercayaan masih diperoleh meskipun ragam tidak diketahui. Maka untuk sampel kecil (n<30) $\sigma_1{}^2$ dan $\sigma_2{}^2 = \sigma^2$ tidak diketahui, harus diduga dengan simpangan baku bersama Sp, yaitu:

$$Sp = \sqrt{\frac{(n_1 - 1)\, s_1{}^2 + (n_2 - 1)\, s_2{}^2}{n^1 + n_2 -^2}}$$

Peubah normal baku T menjadi:

$$\mathrm{T} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{Sp\sqrt{(1/n_1) + (1/n_2)}}$$

Jika $\bar{x}_1$ dan $\bar{x}_2$ merupakan nilai rata-rata contoh acak bebas yang diambil dari dua populasi bebas yang berukuran kecil n_1 dan n_2, maka selang kepercayaan (1-α)100% bagi μ_1 dan μ_2 adalah:

$$(\bar{x}^1 - \bar{x}^2) - t½\alpha\, sp \sqrt{\frac{1}{n^1} + \frac{1}{n^2}} < \mu^1 - \mu^2 < (\bar{x}^1 - \bar{x}^2)$$
$$+ \; t½\alpha\, sp \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad (13.18)$$

Selang kepercayaan (1- α)100% bagi μ_1-μ_2 dua populasi bersifat saling bebas dengan ragam tidak diketahui dan tidak sama $\sigma_1{}^2 \neq \sigma_2{}^2$ adalah:

$$(\bar{x}_1 - \bar{x}_2) - t½\alpha\, sp\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t½\alpha\, sp\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}} \quad (13.19)$$

Derajat bebas (df):

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

Contoh:
Seorang dosen mata kuliah statistik yang mengajar di Fakultas Ekonomi dan Bisins ingin mengetahui apakah ada perbedaan daya Tarik dan pemahaman mahasiswa dan mahasiswi fakultas ekonomi dan bisnis terhadap mata kuliah yang diajarkan. Waktu ujian diberikan waktu yang sama. Setelah ujian diperiksa, dosen tersebut mengambil sampel acak sejumlah 20 nilai yang dipilih.

Mahasiswa FEB: 70,75,81,77,83,68,85,83,88,90
Mahasiswi FEB: 65,75,82,63,79,76,74,81,77,78

Gunakan selang kepercayaan 90%. Diasumsikan bahwa kedua populasi yang menyebar menghampiri normal dengan ragam yang berbeda.

Mahasiswa FEB			**Mahasiswi FEB**		
$\bar{x}_1$	$x_1 - \bar{x}_1$	$(x_1 - \bar{x}_1)^2$	$\bar{x}_2$	$x_2 - \bar{x}_2$	$(x_2 - \bar{x}_2)^2$
70	-10	100	65	-10	100
75	-5	25	75	0	0
81	1	1	82	7	49
77	-3	9	63	-12	144
83	3	9	79	4	16
68	-12	144	76	1	1
85	5	25	74	-1	1
83	3	9	81	6	36

Mahasiswa FEB			Mahasiswi FEB		
$\bar{x}_1$	$x_1 - \bar{x}_1$	$(x_1 - \bar{x}_1)^2$	$\bar{x}_2$	$x_2 - \bar{x}_2$	$(x_2 - \bar{x}_2)^2$
88 90	8 10	64 100	77 78	2 3	4 9
	-	486 s_1^2 = 486/10-1 s_1^2 = 486/9 = 54		-	360 s_2^2 = 360/10-1 s_2^2 = 360/9 = 40

$$\mathrm{df} \cong \frac{\left(\frac{54}{10}+\frac{40}{10}\right)^2}{\frac{\left(\frac{54}{10}\right)^2}{10-1}+\frac{\left(\frac{40}{10}\right)^2}{10-1}} = 17{,}602 \cong 18$$

Pada $\overline{(x_1 - \bar{x}_2)}$ = 80 – 75 = 5

Selang kepercayaan 90% → 10% → 0,05

Untuk df = 18 diperoleh $Z_{0,05}$ = 1,734 selang kepercayaan bagi $\mu_1 - \mu_2$ adalah :

$$(\bar{x}_1 - \bar{x}_2) - t½\alpha\, sp\sqrt{\frac{s_1{}^2}{n_1}+\frac{s_2{}^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t½\alpha\, sp\sqrt{\frac{s_1{}^2}{n_1}+\frac{s_2{}^2}{n_2}}$$

$$5 - 1{,}734\sqrt{\frac{54}{10}+\frac{40}{10}} < \mu_1 - \mu_2 < 5 + 1{,}734\sqrt{\frac{50}{10}+\frac{40}{10}}$$

$$5 - 5{,}32 < \mu^1 - \mu^2 < 5 + 5{,}32$$

$$\text{menjadi} - 0{,}32 < \mu^1 - \mu^2 < 10{,}32$$

selang kepercayaan 90% berapa pada rata-rata nilai mahasiswa dan mahasiswi adalah -0,32 dan 10,32. Karena menunjukan μ_1-μ_2 = 0, maka belum dapat menyimpulkan bahwa kemampuan daya tarik mahasiswa feb lebih baik dibandingkan mahasiswi feb.

13.6 PENDUGAAN INTERVAL PROPORSI

13.6.1 Proporsi Berukuran Besar

p=x/n yaitu proporsi antara banyaknya percobaan berhasil dengan jumlah kalinya percobaan n dilakukan. Jika $\hat{p}$ merupakan penduga bagi parameter p, maka penghampiran distribusi proporsi akan mengambil nilai rata-rata $\mu\hat{p}$ = P dan simpangan baku $\sigma\hat{p} = \sqrt{\mathrm{p}(1-\mathrm{p})/\mathrm{n}}$.. selang kepercayaan bagi proporsi p adalah suatu selang dua nilai, sehingga probabilitas selang tersebut adalah (1-α)100%. Oleh karena itu, jika $\hat{p}$ adalah rata-rata contoh dengan n berukuran cukup besar, maka selang kepercayaan (1- α)100% bagi parameter populasi adalah:

$$\hat{p} - Z½\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + Z½\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Contoh:

Seorang Arsitektur sedang meneliti tentang struktur bangunan rumah tinggal orang jawa yang terkenal klasik dan tradisional di daerah Yogyakarta. Dari penelitian tersebut diambil contoh acak 100 rumah yang diteliti. Setelah dilakukan penelitian terhadap 50 rumah memiliki struktur hunian rumah bergaya rumah orang jawa. Tentukan selang kepercayaan 95% bagi proporsi yang sesungguhnya orang yang memiliki struktur rumah orang jawa.

Jawaban

P = 50/100 = 0,5

Selang kepercayaan 95% atau $Z_{0,025}$ = 1,96.

$$0{,}5 - 1{,}96\sqrt{\frac{0{,}5(1-0{,}5)}{100}} < p < 0{,}5 + 1{,}96\sqrt{\frac{0{,}5(1-0{,}5)}{100}}$$

$$0{,}5 - 1{,}96\,(0{,}0025) < p < 0{,}5 + 1{,}96\,(0{,}0025)$$

$$0{,}4951 < p < 0{,}5049$$

13.6.2 Pendugaan Beda Dua Proporsi

Pendugaan beda dua proporsi diperlukan apabila seorang peneliti menghadapi dua populasi berlainan yang merupakan populasi-populasi yang bebas satu sama lain. Selang kepercayaan bagi p1 – p2 dapat disusun berdasarkan distribusi proporsi penarikan contoh $\hat{p}_1$ dan $\hat{p}_2$ yang menghampiri distribusi normal. Selanjutnya mengganti p_1 dan p_2 dengan nilai dugaan masing-masing $\hat{p}_1$ dan $\hat{p}_2$. Selang kepercayaan beda proporsi bagi p_1 dan p_2 yaitu ukuran penarikan contoh besar selang kepercayaan (1-)100% bagi beda dua parameter populasi, dirumuskan:

$$(\hat{p}_1 - \hat{p}_2) - Z½\alpha \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2)$$

$$+ Z½ \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Contoh:
PT. Oilme memproduksi kecap manis berkemasan botol dan refill .omset penjualan untuk kemasan botol lebih besar 30% dibandingkan penjualan pada kemasan refill. Menurut hasil survey penelitian, ibu rumah tangga menunjukan bahwa 80 konsumen dari 500 konsumen kecap cap oilme menyukai kemasan botol dan 40 diantara 500 konsumen kecap cap oilme menyukai refill. Buatlah selang kepercayaan 90% bagi selisih persentase penjualan kecap cap oilme tersebut. Apakah selisih sebesar 30% dapat dipercaya?

Jawaban
Nilai dugaan titik bagi p

N = 500

$\hat{p}_1$ = 80/500 = 0,16

$\hat{p}_2$ = 40/500 = 0,08

$\hat{p}_1 - \hat{p}_2$ = 0,16 – 0,08 = 0,08

Selang kepercayaan = 90%, $Z_{0,05}$ = 1,64

$$= (\hat{p}_1 - \hat{p}_2) - Z½\alpha \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2)$$

$$+ Z½ \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$= (\hat{p}_1 - \hat{p}_2) - Z½\alpha \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2)$$

$$+ 1,64 \sqrt{\frac{0,16(0,84)}{500} + \frac{0,08(0,92)}{500}}$$

$$= 0,08 - 1,64(0,0203) < p_1 - p_2 < 0,08 + 1,64(0,0203)$$

$$= 0,046 < p_1 - p_2 < 0,113$$

Selang kepercayaan 90% bagi selisih proporsi penjualan kecap kemasan botol dan kemasan refill antara 4,6 – 11,3%. Pernyataan pada persentase penjualan kemasan botol kecap manis cap oilme lebih besar 30% dari kemasan refill dapat dipercaya.

13.7 PENDUGAAN INTERVAL VARIANS DAN SIMPANGAN BAKU

Suatu masalah yang sering dihadapi para pengambil keputusan adalah menyangkut besarnya variabilitas populasi, sehingga perlu dilakukan pembuatan pendugaan ragam populasi. Karena ukuran variabilitas adalah ragam, maka perlu dilakukan pendugaan ragam.

Setiap suku-suku dalam pertidaksamaan dibagi dengan pembilang $(n-1)s^2$. selanjutnya, diatur kembali pertidaksamaan sehingga akan diperoleh:

$$P\left(\frac{(n-1)s^2}{x^2\alpha/_2} < \sigma^2 < \frac{(n-1)s^2}{x^2{}_{1}-\alpha/_2}\right)$$

Jika s^2 merupakan ragam contoh berukuran n yang ditarik dari populasi, maka selang kepercayaan (1 - α)100% bagi ragam populasi σ^2 dinyatakan dalam bentuk selang kepercayaan:

$$\frac{(n-1)s^2}{x^2\alpha/_2} < \sigma^2 < \frac{(n-1)s^2}{x^2{}_1-\alpha/_2}$$

Contoh Soal:

Seorang Ahli Gizi ingin mengetahui proporsi balita laki-laki yang menderita kurang gizi (busung lapar) yang bermukim di wilayah pedalaman. Diambil contoh acak rata-rata n = 40 balita laki-laki dan memperoleh 15 balita laki-laki menderita kurang gizi. Simpangan baku s = 4. Dengan menggunakan selang keyakinan 95%. Dugaan varians dan simpangan baku pada balita laki-laki yang menderita kurang gizi!

Jawab:

n-1 = 15-1 = 14

s = 4

untuk mendapatkan selang kepercayaan 95%, tentukan = 0,05. df = 14 yang diperoleh $_{0,025}$ = 26,119 dan$_{0,975}$ = 5,63 sehingga:

$$= \frac{(n-1)s^2}{x^2\alpha/_2} < \sigma^2 < \frac{(n-1)s^2}{x^2{}_1-\alpha/_2}$$

$$= \frac{(15-1)(4^2)}{x^2 0{,}025} < \sigma^2 < \frac{(15-1)(4^2)}{x^2 0{,}975}$$

$$= \frac{(15-1)(4^2)}{26{,}119} < \sigma^2 < \frac{(15-1)(4^2)}{5{,}63}$$

$$= \frac{224}{26{,}119} < \sigma^2 < \frac{224}{5{,}63}$$

$$=8{,}57 < \sigma^2 < 39{,}78$$

$$=2{,}927 < \sigma < 6{,}31$$

Bab 14

Pengujian Hipotesis Deskriptif

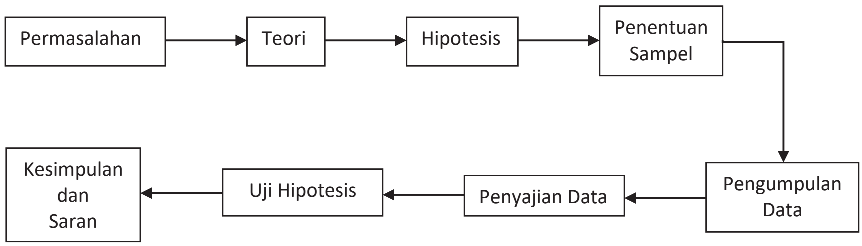
14.1 PENGANTAR

Pengujian hipotesis terdiri dari tiga jenis, yaitu hipotesis deskriptif, komparatif dan asosiatif. Masing-masing ujian mempunyai ketentuan dan syarat yang harus dipenuhi sebelum dapat melakukan analisis. Ketiga pengujian juga membedakan analisis berdasarkan jenis data yang diperoleh baik data nominal, ordinal maupun interval. Oleh karena itu, pada tiga bab terakhir ini fokus utama pembahasan adalah menerangkan penggunaan ketiga pengujian hipotesis tersebut.

14.2 PENELITIAN STATISTIK NONPARAMETRIK

14.2.1 Pengertian Penelitian

Penelitian pada dasarnya merupakan cara ilmiah utuk mendapatkan data dengan tujuan dan kegunaan tertentu. Cara ilmiah berarti kegiatan penelitian itu didasarkan pada ciri-ciri kelimuan, yaitu *rasional, empiris, dan sistematis.* Rasional berarti kegiatan penelitian yang dilakukan dapat diterima oleh akal, sehingga terjangkau oleh nalar manusia. Empiris berarti cara-cara yang dilakukan dengan melalui pengamatan, sehingga orang lain dapat menilai, melihat dan mengetahui langkah-langkah yang digunakan dalam penelitian tersebut. Sistematis artinya setiap proses selama penelitian berlangsung menggunakan cara yang logis dan terukur. Berdasarkan definisi ketiga unsur penelitian ini dapat ditunjukkan melalui gambar 14.1 berikut:



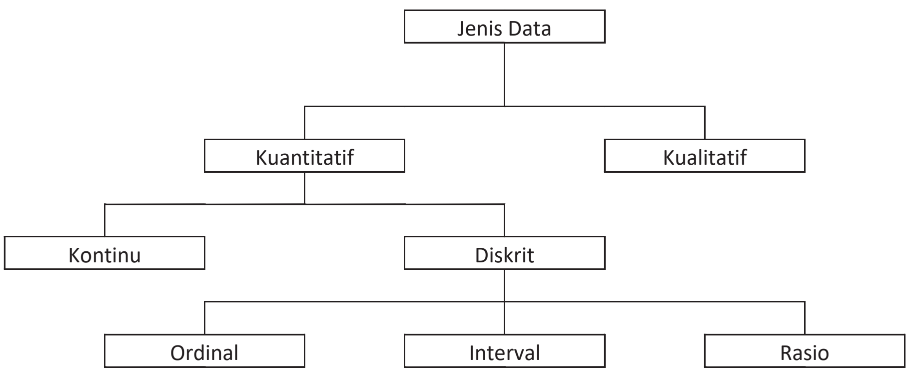
GAMBAR 14.1 ***Alur Penelitian***

Data yang diperoleh melalui penelitian itu mempunyai kriteria tertentu yang valid, yang berarti tingkat ketepatan dari seluruh data dalam penelitian adalah riil dan sedang terjadi pada objek data yang sedang diteliti. Selama proses penelitian, data yang valid dapat diketahui dengan menggunakan uji realibitas dan objektivitas. Reliabilitas dapat menunjukkan konsistensi data dalam interval waktu tertenu. Sedangkan, objektivitas erat kaitannya dengan hal-hal yang dinilai *Common* atau dianggap biasa oleh orang-orang. Contoh, dari 40 pelajar SMA kelas X yang diwawancarai, sebanyak x pelajar berpendapat mata pelajaran matematika itu susah sedangkan mata pelajaran olahraga itu menyenangkan. Selain itu, data yang reliabel belum tentu valid.

Contoh seorang pegawai mengatakan setiap hari ada rapat, faktanya pegawai tersebut mempunyai jadwal 3 kali dalam seminggu. Objektivitas data juga tidak akan dinilai valid apabila jumlah persentase benar lebih kecil dari yang salah. Contohnya, dari hasil survei yang dilakukan, sejumlah 94% orang berpandangan bahwa bumi datar, sedangkan hanya 6% yang berpendapat bumi bulat. Setiap penelitian mempunyai kegunaan dan tujuan tertentu. Terdapat tiga macam sifat dalam penelitian yaitu penelitian bersifat penemuan baru, pembuktian, dan pengembangan. Penelitian untuk penemuan berarti data yang diperoleh dan diolah merupakan data baru yang belum pernah dipublikasi atau diketahui sebelumnya. Penelitain berupa pembuktian berati data yang diperoleh berasal dari lembaga penerbit data tersebut atau bersifat data sekunder. Dalam praktiknya, penelitian pembuktian dilakukan untuk menunjukkan kebenaran dari sebuah teori. Penelitian terakhir yaitu penelitian pengembangan berarti untuk memperdalam dan memperkuat ilmu pengetahuan serta teori-teori yang telah ada. Sehingga, berdasarkan hasil penelitian ini nantinya, manusia dapat memahami, memecahkan atau mengantisipasi masalah di masa depan.

14.2.2 Jenis-Jenis data Penelitian

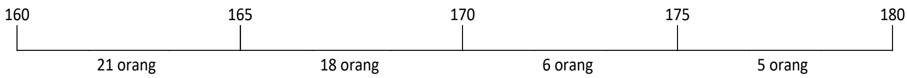
Seperti yang sudah dijelaskan pada bab awal dalam buku ini, sumber data terdiri dari 3 yaitu data primer, sekunder dan tersier. Setelah mengetahui sumber data tersebut, kita dapat melihat jenis data yang terkumpul. Tujuannya adalah untuk melihat apakah data yang tersedia masuk ke dalam kelompok data kualitatif atau kuantitatif. Hal ini bertujuan untuk memudahkan dalam pembentukan hipotesis, seperti yang ditunjukkan dalam gambar 14.2 berikut:



GAMBAR 14.2 ***Jenis-Jenis Data dalam Penelitian***

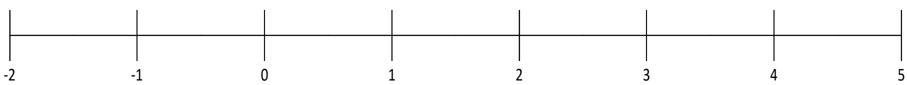
Data dalam penelitian terdiri dari dua jenis, kualitatif dan kuantitatif. Data kualitatif umumnya berbentuk perkataan, kalimat, atau gambar. Sedangkan data kuantitatif adalah data dalam bentuk angka atau data kualitatif dengan penggunaan skala/skoring.

Seperti yang ditunjukkan dalam Gambar 14.2, data kuantitatif terdiri dari dua yaitu diskrit dan kontinu. Data diskrit merupakan hasil perhitungan atas beberapa kelompok, misalnya penghuni rumah sewa kos pendopo berjumlah 50 orang laki-laki, terdiri atas 25 orang Jawa, 10 orang Sunda, 5 orang palembang, 4 orang Batak, 1 orang Aceh dan 5 orang Padang. Data kontinu pula adalah data yang mempunyai variasi yang disusun berdasarkan tingkat dari hasil pengukuran. Dalam kelompok data kontinu terdiri dari tiga jenis yaitu Ordinal, Interval dan Rasio. Contohnya, dari 50 orang laki-laki penghuni rumah sewa kos pendopo yang mempunyai tinggi badan rentang 160-165 berjumlah 18 orang, tinggi badan 166-170 berjumlah 21, tinggi badang 170-175 berjumlah 6 orang dan sisanya 175-180 berjumlah 5 orang. Lihat Gambar 14.3 berikut:



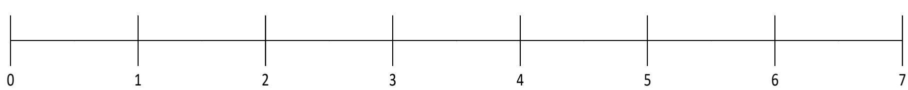
GAMBAR 14.3 ***Data Ordinal***

Data interval adalah data dengan mempunyai jarak sama tetapi tidak mempunyai nilai absolut/mutlak. Salah satu data yang tergolong dalam kelompok data interval adalah data likert. Data ini dapat dibuat dengan mengelompokkan data dalam bentuk peringkat seperti pada Gambar 14.4 di bawah ini.



GAMBAR 14.4 ***Data Interval***

Data rasio pada umumnya dalam bentuk persentase dengan nilai jarak yang sama. Contohnya adalah berat, panjang, volume, dan bobot. Perbedaan antara data ordinal, interval dan rasio terletak pada nila hitung penjumlahan atau perkalian dalam aljabarnya. Contoh, air dengan suhu 24^{0}C dicampurkan ke dalam air dengan suhu 14^{0}C tidak akan menambah suhu menjadi 38^{0}C tetapi menjadi berkisar antara 16^{0}C-20^{0}C. Oleh karena itu, data rasio adalah data yang sering digunakan untuk menggambarkan fenomena kejadian dan merupakan data yang paling teliti.



GAMBAR 14.5 ***Data Rasio***

14.2.3 Pembentukan Hipotesis

Hipotesis merupakan jawaban sementara yang harus ditentukan dalam rumusan masalah penelitian. Hipotesis dapat diuji melalui beberapa cara dan didasarkan pada teori yang digunakan. Dinilai jawaban sementara karena belum dibuktikan dan memperoleh fakta sebenarnya. Dalam penelitian yang menggunakan analisis statistik inferensial, terdapat dua hipotesis yang perlu diuji, yaitu hipotesis penelitian dan hipotesis statistik. Hipotesis penelitian dilakukan untuk membuktikan apakah hipotesis yang diajukan terjadi atau tidak. hipotesis statistik dilakukan untuk melihat apakah hipotesis penelitian dapat dibuktikan atau tidak dengan menggunakan data sampel. Hipotesis statistik dinyatakan dalam bentuk H_0 (hipotesis nol) dan H_1 (hipotesis alternatif). Pernyataan yang hendak dibuktikan oleh peneliti, yaitu hipotesis peneliti dikonversikan menjadi H_1. Hipotesis penelitian dinyatakan (atau dianggap) terbukti jika uji hipotesis menghasilkan penolakan H_0.

Menurut tingkat kejelasan (Level of explanation) peubah yang diteliti, terdapat tiga bentuk hipotesis yang dapat dirumuskan dan diuji, yaitu:

1. Hipotesis Deskriptif

   Hipotesis ini merupakan pendugaan terhadap nilai peubah dalam satu sampel walaupun di dalamnya terdapat beberapa kategori.

   Contoh:

   Ho: kecenderungan mahasiswa memilih mata kuliah yang mudah

   H1: kecenderungan mahasiswa memilih mata kuliah berdasarkan dosen pengajar

2. Hipotesis Komparatif

   Hipotesis ini merupakan pendugaan terhadap perbandingan pada dua sampel atau lebih. Komparasi yang dilakukan didasarkan pada hal yang berkaitan, seperti:

   a. komparasi berpasangan dalam dua sampel dan lebih dari dua sampel
   b. komparasi independen dan dua sampel dan lebih dari dua sampel

Contoh sampel berpasangan untuk komparasi dua sampel

Ho : Tidak terdapat perbedaan nilai penjualan sebelum dan sesudah adanya iklan

H1 : Terdapat perbedaan nilai penjualan sebelum dan sesudah adanya iklan

Contoh sampel independen untuk komparasi tiga sampel

Ho : Tidak adanya perbedaan antara mahasiswa, pegawai, dan pengusaha dalam berbelanja melalui *e-commerce*

H1 : Terdapat perbedaan dikalangan mahasiswa, pegawai, dan pengusaha dalam berbelanja melalui *e-commerce*

3. Hipotesis Asosiatif

   Hipotesis ini merupakan pendugaan untuk melihat hubungan antara dua peubah atau lebih.

   Contoh:

   Ho : Tidak ditemukannya hubungan antara usia dengan peminatan dalam bermain game online

   H1 : Ditemukannya hubungan antara usia dengan peminatan dalam bermain game online

Contoh:

Berdasarkan hasil nilai rapor Siswa SMP kelas IX diperoleh data banyak yang mendapatkan nilai di bawah standard kelulusan. Saat proses wawancara, peneliti mendapati metode pengajaran yang diberikan dirasa kurang memuaskan, akibatnya pelajar susah memahami materi dan perlu adanya pengembangan konsep pengajran untuk mata pelajaran tersebut. Sebelum digunakan, metode baru ini harus diuji dalam penelitian untuk melihat kebenaran apakah metode pengajaran yang baru lebih baik berbanding metode pengajaran yang baru.

Hipotesis penelitian ini adalah proses pengajaran dengan menggunakan metode baru untuk siswa SMP kelas X lebih baik berbanding metode pengajaran lama.

Hipotesis statistik ini adalah:

H_0: $\mu_1 \leq \mu_2$

H_1: $\mu_1 > \mu_2$

$\mu 1$: rerata nilai tes matematika siswa SMU yang memperoleh pengajaran dengan metode baru

$\mu 2$: rerata nilai tes matematika siswa SMU yang memperoleh pengajaran

Hipotesis penelitian dinyatakan (dianggap) terbukti benar apabila pengajaran matematika pada siswa SMP dengan metode baru memberi hasil yang lebih baik daripada dengan metode yang lama jika analisis statistik terhadap data sampel menghasilkan penolakan hipotesis nol. Dalam keadaan tertentu yang relatif lebih jarang ditemukan, yaitu pada 'penelitian negatif' peneliti semata-mata bertujuan membuktikan bahwa sesuatu yang baru adalah 'sama baiknya' ataupun 'tidak lebih buruk' daripada hal yang sama. Hipotesis statistiknya adalah:

$H_0 : \mu_1 = \mu_2 \quad vs \quad H_1 : \mu_1 \neq \mu_2$

atau:

$H_0 : \mu_1 \geq \mu_2 \quad vs \quad H_1 : \mu_1 < \mu_2$

μ_1 : rerata 'sesuatu' yang baru

μ_2 : rerata 'hal' yang lama

Di sini tujuan peneliti dianggap tercapai jika hipotesisnya nol tidak ditolak, dengan kata lain hipotesis nol 'dianggap' benar.

14.2.4 Langkah-Langkah Penentuan Pengujian Hipotesis

Dalam ilmu statistik terdapat dua macam statistik inferensial yang dapat digunakan untuk menguji hipotesis, yaitu statistik parametrik dan statistik non-parametrik. Keduanya dikerjakan dengan menggunakan data sampel yang berasal dari sumber data random. Statistik parametrik cenderung digunakan pada data interval dan rasio dengan syarat data yang diambil berdistribusi normal. Statistik nonparametrik digunakan untuk data nominal/diskrit dan ordinal namun tidak perlu syarat data harus berdistribusi normal seperti pada statistik parametrik.

Penentuan teknik statistik yang sesuai untuk pengujian hipotesis dalam penelitian dapat dilakukan dengan melihat data yang tersedia apakah dalam bentuk ordinal atau nomial serta apakah hipotesis yang dibangun adalah hipotesis deskriptif, komparatif atau asosiatif. Pada tabel 14.1 berikut akan menunjukkan teknik statistik yang sebaiknya digunakan untuk pengujian hipotesis.

TABEL 14.1 ***Panduan Pemilihan Statistik Nonparametrik untuk Pengujian Hipotesis***

Bentuk Data	Hipotesis					
	Deskriptif (satu peubah)	komparatif (dua sampel)		komparatif (> dua sampel)		Asosiatif
		Related	Independen	Related	Independen	
Nominal	Binomial	Mc Nemar	Fisher Exact Probability	$\chi 2$ for k sample	$\chi 2$ for k sample	Contingency Coefficient C
	$\chi 2$ One Sample		$\chi 2$ Two Sample	Cochran Q		
Ordinal	Run Test	Sign test	Median test	Friedman Two Way-Anova	Median Extension	Spearman Rank Correlation
		Wilcoxon matched parts	Mann-Whitney U test		Kruskal-Wallis One Way Anova	Kendall Tau
			Kolmogorov Smirnov			
			Wald-Woldfowitz			
Interval Rasio	T test	T-test of* Related	T-test of* independent	One-Way Anova*	One-Way Anova*	Pearson Product Moment *
				Two Way Anova*	Two Way Anova*	Partial Correlation*
						Multiple Correlation*

Langkah-langkah Umum pada Uji Hipotesis:

1. Tentukan jenis uji/analisis statistik yang akan digunakan (sesuaikan dengan jenis data skala pengukurannya).
2. Rumuskan hipotesis nol dan hipotesis alternatif.
3. Tentukan tingkat signifikansi a, kekuatan uji $(1-\beta)$, selisih parameter minimum yang hendak dideteksi.
4. Hitung ukuran sampel minimum yang dibutuhkan. Jika perlu lakukan studi pendahuluan untuk memperoleh data yang dibutuhkan namun belum ada untuk perhitungan ukuran sampel.

5. Pilih statistik penguji yang akan digunakan (sesuai dengan uji/analisis).
6. Tentukan sampel dan lakukan pengumpulan data.
7. Hitung nilai statistik penguji dari data sampel. Periksa apakah nilai ini terletak pada daerah kritis atau tidak.
8. Hitung nilai p dan buat kesimpulan apakah H_0 ditolak atau tidak ditolak.
9. Hitung dan sajikan interval konfidensi 100 $(1-a)$% untuk parameter yang dikaji.
10. Jika H_0 tidak ditolak dan ukuran sampel yang diperoleh tidak mencukupi ukuran minimum yang dibutuhkan, hitung kembali kekuatan uji yang sesungguhnya.

Penentuan kekuatan uji $(1-\beta)$ dan selisih parameter minimum yang hendak dideteksi pada langkah ke-3 serta langkah-langkah ke-4, 6, 9, dan 10 tidak termasuk dalam uji hipotesis yang sebenarnya, namun dalam praktik harus dikerjakan harus dikerjakan dalam urutan tersebut sebagai langkahlangkah metode penelitian.

14.3 PENGUJIAN HIPOTESIS DESKRIPTIF (SATU SAMPEL)

Hipotesis deskriptif merupakan salah satu pendugaan terhadap nilai peubah dalam satu sampel. Pada data nominal dalam statistik nonparametrik untuk menguji hipotesis deskriptifnya dapat menggunakan uji Test Binomial dan Chi Kuadrat ($\chi 2$) sedangkan apabila data dalam bentuk ordinal maka uji yang digunakan adalah Run Test.

14.3.1 Test Binomial

Pengujian hipotesis menggunakan test binomial dalam satu kelompok data/populasi terdiri dari dua kelas, pertama data berbentuk nominal dan jumlah sampelnya adalah kecil (<25). Contohnya, kelas jenis kelamin, laki-laki dan perempuan, pendapatan seperti berpenghasilan tinggi dan rendah, prestasi, senior dan junior dan lain sebagainya. Selanjutnya, nilai populasi ini akan diteliti menggunakan sampel yang dari jumlah populasi tersebut.

Data sampel yang telah diperoleh dari jumlah populasi data makan peneliti dapat menguji hipotesis statistiknya. Biasanya pengujian yang dilakukan adalah untuk melihat apakah ada perbedaan antara data dalam populasi dengan data yang terdapat pada sampel. Uji yang dilakukan untuk melihat perbedaan ini menggunakan Test Binomial. Jadi, secara definisi Test Binomial adalah test untuk menguji hipotesis deskriptif (satu sampel) dengan jenis data nominal yang mempunyai dua kelas. Oleh karena itu, uji test binomial sangat sesuai untuk pengujian hipotesis dengan ukuran sampel yang kecil, sedangkan uji chi square sangat tidak mendukung.

Test binomial merupakan test yang dimana distribusi datanya berbentuk binomial. Distribusi binomial mempunyai ciri yaitu terdiri dari dua kelas. Satu populasi berjumlah N mempunyai satu kelas x dan kelas lainnya adalah N-x. Peluang untuk nilai x dalam satu kategori dan N-x dapat dirumus sebagai berikut:

$$P^{(x)} = \binom{N}{x} P^{x} Q^{N \,.\, x}$$

Di mana P adalah proporsi kasus yang diharapkan dalam salah satu kategori dan Q adalah kategori lainnya. Besarnya nilai Q dapat ditentukan dengan Q = 1 – P. Sedangkan, nilai pada rumus dapat dihitung dengan menggunakan rumus berikut:

$$\binom{N}{x} = \frac{N!}{x!\ (N-x)!} P^{x} Q^{N \,.\, x}$$

Nilai N adalah nilai faktorial di mana *N*! = N(N-1)(N-2)....dst

Contoh: nilai dari 4! Adalah?

4!= 4(4-3) (4-2) (4-1)

4!= 4(1)(2)(3)

5!= 24

Pada lampiran terakhir Tabel V menunjukkan harga faktorial bagi nilai $\frac{N}{x}$ sampai dengan 20 dan Tabel IV menunjukkan nilai koefisien binomial untuk harga N = 1 s/d 25.

Pengaplikasian uji binomial dalam statistik cenderung lebih sederhana. Penentuan Ho diterima atau ditolak dapat dinilai berdasarkan dengan nilai perbandingan nilai P dalam tabel dengan derajat kesalahan yang ditetapkan. Apabila dalam sampel terdiri dari 20 hasil pengamatan dan sampel tersebut adalah 4. Berdasarkan Tabel IV nilai dari harga P adalah 0.006. Dengan derajat kesalahan pada $\alpha = 0.01$ maka keputusan yang dapat diambil dari hasil pengujian hipotesis tersebut adalah Ho diterima dan H1 ditolak karena nilai P lebih kecil dari α. Ketentuannya adalah apabila nilai $p \geq \alpha$ maka Ho ditolak.

Contoh:
Pabrik mobil berhasil memproduksi dua jenis mobil khusus keluarga dengan bahan bahar bensin dan solar. Perusahaan tersebut ingin mengetahui apakah keluarga yang berminat membeli mobil mempertimbangkan mobil dengan kedua bahan bakar. Dari hasil pengumpulan data diperoleh sampel berjumlah 18 keluarga yang terkumpul dari seluruh populasi penjualan mobil. Dipilih secara acak didapati 10 keluarga memilih membeli mobil berbahan bakar solar dan 8 keluarga berbahan bakar bensin. Berdasarkan kasus di atas, tentukan keputusan apa yang sebaiknya diambil oleh perusahaan.

Jawaban:

1. Dari kasus di atas, judul yang tepat adalah kecenderungan keluarga dalam memilih mobil berdasarkan bahan bakar
2. Peubah penelitian adalah jenis mobil
3. Rumusan masalah dalam penelitian tersebut adalah bagaimana kecenderungan keluarga membeli mobil? Apakah cenderung memilih mobil dengan penggunaan bahan bakar solar atau bensin?
4. Hipotesis yang diajukan adalah:
   Ho : Jumlah keluarga membeli mobil berbahan bakar bensin dan solar sama/tidak berbeda
   H1 : Jumlah keluarga membeli mobil berbahan bakar bensin dan solar berbeda

   Dari kasus ini didapati:
   Ho : $p_1 = p_2 = 0,5$
   H1 : $p_1 \neq p_2 \neq 0,5$
5. Sampe sebagai sumber data untuk pengujian hipotesis adalah:
   Sebagian dari kelompok keluarga yang dipilih secara acak. Dalam sampel terdapat dua kelompok keluarga yang membeli mobil dengan bahan bakar

solar 10 keluarga dan keluarga yang membeli mobil berbahan bakar bensin sebanyak 8 keluarga.

6. Teknik pengumpulan data yang digunakan adalah pengamatan saat keluarga membeli mobil di *showroom* mobil.
7. Hasil penelitian dapat ditunjukkan pada tabel berikut:

TABEL 14.2 *Kecenderungan Keluarga membeli mobil*

Bahan Bakar Mobil	Frekuensi Pembeli
Mobil berbahan Bakar Solar	10
Mobil berbahan Bakar Bensin	8
Jumlah	18

8. berdasarkan hipotesis yang dirumuskan yaitu hipotesis deskriptif (satu sampel) dengan data berbentuk nominal, dan dengan jumlah anggota sampel <25. Oleh karena itu, dipertimbangkan teknik statistik yang sesuai untuk penelitian ini adalah pengujian dengan test binomial.

   Kasus di atas menjabarkan bahwa jumlah sampel dalam penelitian berjumlah 18 keluarga dengan (N)=18. Hasil pendataan mendapati keluarga yang membeli mobil berbahan bakar solar berjumlah 10 sedangkan berbahan bakar bensin berjumlah 8, berarti nilai frekuensi terkecil dari sampel ini adalah 8, dengan (x)= 8. Merujuk pada tabel angka binomial dengan syarat N=18 dan x=8, maka nilai koefisien binomialnya sebesar = 0,407. Bila taraf kesalahan α adalah 1% yang berarti = 0,01 maka harga p sebesar 0,271 ternyata lebih besar dari 0,01 (0,407 > 0,01). Harga p > harga α maka Ho diterima dan H1 ditolak. Kesimpulannya adalah frekuensi keluarga memilih dua jenis mobil sama yaitu 50% membeli mobil berbahan bakar solar dan membeli mobil berbahan kasar bensin 50%.
9. Kesimpulan: ada uda kecenderungan yang sama di keluarga dalam membeli mobil yaitu mobil berbahan bakar bensin dan solar.
10. Saran yang diberikan dalam penelitian ini adalah untuk perusahaan produsen mobil sebaiknya memperoduksi mobil dalam jumlah yang sama.

14.3.2 Chi Kuadrat (χ^2)

Chi kuadrat untuk sampel merupakan pengujian hipotesis deskriptif apabila populasi mempunyai dua kategori kelas atau lebih serta data yang diperoleh berbentuk nominal dengan jumlah sampel yang besar. Hipotesis dekspritif ini adalah pendugaan atau estimasi terhadap perbedaan frekuensi antara kategori satu dan kategori lain dalam sebuah sampel tentang sesuatu hal.

Rumus untuk perhitungan Chi Kuadrat adalah seperti berikut:

$$X^2 = \sum_{i=1}^{k} \frac{(fo - fh)^2}{fn}$$

Dimana:

χ^2 = Chi Kuadrat
fo = Frekuensi Observasi
fh = Frekuensi Diharapkan

Contoh:

Dalam sebuah organisasi kampus sedang melaksanakan penelitian yang bertujuan untuk melihat peluang perempuan terpilih sebagai ketua organisasi. Ditetapkan dalam penelitian ini bahwa populasinya adalah seluruh mahasiswa aktif. Dari jumlah keseluruhan mahasiswa di kampus, terpilih sampel secara acak berjumlah 300 mahasiswa yang berpartisipasi dalam pemilihan ketua organisasi untuk laki-laki dan perempuan. Berdasarkan perolehan sampel didapati bahwa pemilih perempuan berjumlah 100 mahasiswa dan pemilih laki-laki berjumlah 200 mahasiswa. Tentukan kesimpulan penelitian ini.

Jawaban:

1. Dari kasus di atas, judul yang tepat adalah peluang terpilihnya ketua organisasi laki-laki dan perempuan
2. Peubah penelitian adalah ketua organisasi
3. Penelitian ini meneliti kecenderungan mahasiswa dalam memilih ketua organisasi dengan jumlah sampel sebanyak 300 mahasiswa dengan jumlah pemilih perempuan berjumlah 100 mahasiswa dan pemilih laki-laki bejumlah 200 mahasiswa.

4. Hipotesis yang diajukan adalah:
   Ho: Jumlah mahasiswa yang memilih calon ketua organisasi laki-laki dan perempuan adalah tidak berbeda, artinya baik laki-laki maupun perempuan berpeluang menjadi ketua organisasi
   H1: Jumlah mahasiswa yang memilih calon ketua organisasi laki-laki dan perempuan adalah berbeda, artinya baik laki-laki maupun perempuan mempunyai peluang yang lebih besar menjadi ketua organisasi

   Dari kasus ini didapati:
   Ho: $p_1 = p_2 = 0{,}5$
   H1: $p_1 \neq p_2 \neq 0{,}5$
5. Tempat penelitian dilaksanakan di Kampus XYZ
6. Ketentuan untuk hasil penelitian dalam metode *Chi Kuadrat* apabila nilai X < dari nilai tabel Chi Kuadrat pada tingkat kesalahan tertentu, maka Ho diterima dan H1 ditolak. Tetapi sebaliknya, bila nilai tabel Chi Kuadrat > atau sama dengan nilai tabel maka H1 diterima.
7. Hasil pengumpulan data penelitian

TABEL 14.3 ***Kecenderungan Keluarga membeli mobil***

Calon Ketua Organisasi	Frekuensi diperoleh	Frekuensi diharapkan
Laki-laki	200	150
Perempuan	100	150
Jumlah	300	300

8. Pengujian hipotesis dapat dilakukan dengan mengaplikasikan rumus yang telah dijelaskan sebelumnya. Untuk menghitung nilai Chi Kuadrat menggunakan rumus di atas, maka sebaiknya digunakan tabel pembantu untuk memudahkan pemahaman anda.

Alternatif pilihan	fo	fh	fo - fh	$(fo - fh)^2$	$\frac{(fo - fh)^2}{fh}$
Laki-laki	200	150	50	2500	16,67
Perempuan	100	150	-50	2500	16,67
Jumlah	300	300	0	5000	33,33

Catatan: Sesuai dengan persyaratan penelitian yang telah dibentuk diawal maka nila yang diharapkan untuk kesempatan terpilihnya calon laki-laki dan perempuan adalah sebesar 50 persent. Sehingga, 300 x pemilihan 50% = 150

Merujuk kepada perhitungan tabel di atas, maka nilai Chi kuadrat yang diperoleh adalah sebsar 33,33. Sebelum mengambil keputusan dalam peneliitian maka peneliti harus melihat kepada nilai tabel chi kuadrat yang terdapat di dalam tabel nilai kesalahan. Apabila nilai Chi Kuadrat lebih kecil dari nilai tabel maka Ho diterima, sedangakn apabila lebih besar maka Ho ditolak.

Dalam kasus ini, nilai fo sama dengan nilai fh, maka kita tidak mempunyai nilai derajat kebebasan lainnya selain 1. Sehingga, apabila dk=1 dengan tingkat kesalahan sebesar 5% diperoleh nilai chi kuadrat untuk tabel adalah sebesar 3,841. Kemudian, kita dapat membuat perbandingan berdasarkan kedua nilai ini (nilai chi kuadrat 33,33 > nilai tabel 3,841). Sesuai dengan ketentuan maka Ho ditolak dan H1 diterima. Jadi kesimpulannya adalah peluan laki-laki dan perempuan terpilih menjadi ketua organisasi mahasiswa adalah berbeda. Artinya, mahasiswa cenderung memilih ketua laki-laki berbanding ketua perempuan.

Contoh:
Perusahaan sepatu ingin mengetahui produk mana yang paling diminati oleh pengguna. Oleh karena itu, mereka membuat penelitian dengan melihat perilaku pembeli dengan mengamati tipe sepatu yang paling sering dibeli. Tipe sepatu tersebut dibagi menjadi 4 kelompok yaitu sepatu A, B, C dan D. Dalam pengamatan diperoleh data sampel untuk masing-masing tipe sepatu. Sampel kelompok sepatu A berjumlah 1000 orang, sepatu B berjumlah 900, sepatu C berjumlah 600 dan sepatu D berjumlah 500. Berdasarkan informasi di atas, tentukan kesimpulan apa yang sebaiknya diambil oleh produsen sepatu tersebut.

Jawaban
1. Dari kasus di atas, judul yang tepat adalah peminatan pembeli terhadap tipe sepatu
2. Peubah penelitian adalah tipe sepatu
3. Penelitian ini meneliti kecenderungan pembeli dalam membeli sepatu. Sampel yang berhasil dikumpul selama proses pengumpulan data adalah sebagai berikut; peminat sepatu A berjumlah 1000 orang, sepatu B

berjumlah 900 orang, sepatu C sebanyak 600 orang, sedangkan sepatu D sebanyak 500 orang.

4. Hipotesis yang diajukan adalah:

   Ho: Jumlah pembeli yang membeli sepatu mengikut tipe sepatu tidak berbeda, artinya peluang pembeli masing-masing tipe sepatu adalah sama

   H1: Jumlah pembeli yang membeli sepatu mengikut tipe sepatu berbeda, artinya peluang pembeli masing-masing tipe sepatu adalah beragam

   Dari kasus ini didapati:

   Ho: $p_1 = p_2 = 0{,}5$

   H1: $p_1 \neq p_2 \neq 0{,}5$

5. Tempat penelitian dilaksanakan di toko sepatu merk terkenal di pasar raya
6. Ketentuan untuk hasil penelitian dalam metode *Chi Kuadrat* apabila nilai X < dari nilai tabel Chi Kuadrat pada tingkat kesalahan tertentu, maka Ho diterima dan H1 ditolak. Tetapi sebaliknya, bila nilai tabel Chi Kuadrat > atau sama dengan nilai tabel maka H1 diterima.
7. Hasil pengumpulan data dan perhitungan hipotesis penelitian

TABEL 14.4 ***Kecenderungan pembeli melihat tipe dalam membeli sepatu***

Tipe Sepatu	fo	fh	fo-fh	$(fo-fh)^2$	$\frac{(fo - fh)^2}{fh}$
A	1000	750	250	62.500	83,33
B	900	750	250	22.500	30.000
C	600	750	-150	22.500	30.000
D	500	750	-250	62.600	83,33
Jumlah	3000	3000	0	5000	226,67

Catatan: berdasarkan hipotesis yang telah dibentuk maka untuk nilai fh berdasarkan pada pembagian masing-masing tipe sepatu dengan sama rata.

8. Pengujian hipotesis dilakukan dengan menghitung nilai dalam tabel 14.4 dan dimasukkan ke dalam rumus. Dari hasil perhitungan diperoleh nilai chi kuadrat (dalam tabel 14.4) adalah sebesar 226,67. Derajat kebebasan (dk) untuk kategori kelompok sepatu yang terdiri dari 4 tipe adalah,dk=

n-1= 4-1=3. Dengan diketahui nilai dk adalah sebesar 3 dan tingkat kesalahan sebesar 5% maka diperoleh nilai chi kuadrat dalam tabel adalah 7,815. Dari hasil tersebut maka Chi kuadrat masih lebih besar dari nilai chi kuadrat dalam tabel (22,67 > 7,815). Sehingga, Ho ditolak dan H1 diterima.

9. Kesimpulannya adalah pembeli cenderung memilih sepatu dengan tipe berbeda. Dari data yang dikumpulkan kebanyakan pembeli minat kepada sepatu bertipe A.
10. Saran yang dapat diberikan kepada produsen sepatu tersebut adalah karena tipe sepatu yang paling diminati adalah sepatu bertipe A maka sepatu yang sebaiknya paling banyak diproduksi dan diedarkan adalah sepatu dengan tipe A.

14.3.3 Run Test

Selain kedua test sebelumnya, test binomial dan chi kuadrat, uji run test juga dapat digunakan untuk pengujian hiptesis deskriptif dengan kategori satu sample dan data berbentuk ordinal. Uji ini dilakukan untuk mengukur kerandoman populasi berdasarkan data hasil pengamatan melalui data sampel. Kemudian, dari data tersebut dapat diukur banyaknya nilai "run" yang terjadi. Contoh sederhana dalam uji run test adalah pada uang koin. Koin mempunyai dua mata sisi, gambar (G) dan angka (A). Setelah dilakukan pelemparan sebanyak lima belas kali data yang berhasil dikumpul adalah sebagai berikut:

<u>GGG</u>	<u>AAA</u>	<u>G</u>	<u>AAAA</u>	<u>GG</u>	<u>A</u>	<u>G</u>
1	2	3	4	5	6	7

Pelemparan koin dilakukan sebanyak 7 kali dengan hasil pada lemparan pertama didapati G sebanyak tiga kali, lemparan kedua diperoleh A sebanyak 3 kali, kemudian lemparan ke empat adalah G satu kali, keempat A, kelima G, keenam A dan ketujuh G. Dalam statistik data hasil lemparan koin ini berbentuk ordinal. Sehingga, setelah seluruh data terkumpul maka nilai Ho dapat membandingkan nilai run dalam observasi dengan nilai pada tabel untuk test Run pada persentase signifikan tertentu. Syarat yang harus dipenuhi untuk menilai observasi dari pengukuran nilai test Run adalah apabila nilai tersebut terletak diantara nilai tabel yang kecil dan besar maka Ho diterima dan H1 ditolak.

Perhitungan di atas dilakukan apabila sampel dalam penelitian adalah kecil. Sedangkan untuk sampel dalam jumlah besar maka rumus z yang dapat digunakan adalah sebagai berikut:

$$z = \frac{r - \mu_r}{\sigma_r} = \frac{r - \left(\frac{2n_1n_2}{n_1 + n_2} + 1\right) - 0{,}5}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2}{(n_1 + n_2)^2\,(n_1 + n_2 - 1)}}}$$

Kemudian, untuk nilai (mean) atau atau μ_r dan simpangan baku (σ_r) (dapat dihitung dengan menggunakan dua rumus berikut:

$$\mu_r = \left(\frac{2n_1n_2}{n_1 + n_2} + 1\right) - 0{,}5$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2}{(n_1 + n_2)^2\,(n_1 + n_2 - 1)}}$$

Contoh kasus apabila sampel dalam jumlah kecil:

Dalam penelitian yang sedang dilakukan oleh pengamat ingin melihat pola makan siang di kantin sekolah. Dari data yang terkumpul diperoleh sekelompok pelajar sebanyak 24 orang secara acak untuk diwawancarai tentang waktu terbaik untuk belajar. Pertanyaan yang diutarakan terdiri dari dua jawaban alternatif yaitu apakah setelah pulang sekolah atau malam hari. Wawancarai dilakukan secara berurutan dimulai dari urutan No. 1 hingga 24. Tentukan hasil hipotesis penelitian di atas.

Jawaban

1. Judul penelitian yang sesuai untuk kasus di atas adalah Waktu Terbaik bagi Siswa untuk Belajar
2. Peubah yang diteliti adalah waktu yang dipilih oleh siswa
3. Sampel penelitian ini berjumlah 24 orang, N=24 dengan n1=12 dan n2=12
4. Tempat penelitian dilakukan di kantin sekolah SMA Jaya

5. Hipotesis Penelitian adalah sebagai berikut:

   Ho: Peluang siswa memilih waktu belajar terbaik untuk belajar adalah sama yaitu 50 persen untuk waktu belajar setelah pulang sekolah dan 50 persen untuk malam hari dan urutan wawancara bersifat random

   H1: Peluang siswa memilih waktu belajar terbaik untuk belajar adalah sama yaitu 50 persen untuk waktu belajar setelah pulang sekolah dan 50 persen untuk malam hari urutan wawancara tidak bersifat random
6. Keputusan Uji hipotesis tersebut adalah apabila nilai run observasi berada di antara tabel run kecil dan besar makan Ho diterima dan H1 ditolak.
7. Penyajian Data dilakukan untuk membuktikan hipotesis. Dari hasil wawancara disusun kembali berdasarkan urutan dengan P berarti pulang sekolah dan M untuk malam hari seperti yang ditunjukkan dalam tabel 14.5

TABEL 14.5 ***Jawaban Responden dalam Wawancara***

No.	Jawaban	No.	Jawaban
1.	P	13.	M
2.	P	14.	P
3.	M	15.	P
4.	P	16.	M
5.	M	17.	P
6.	P	18.	M
7.	M	19.	M
8.	M	20.	P
9.	P	21.	M
10.	P	22.	M
11.	M	23.	P
12	M	24.	P
Jumlah Run	8		8

   Berdasarkan hasil wawancara di atas, dapat disimpulkan bahwa nilai hitung untuk test run tersebut adalah (r)=16. Cara menghitung nilai run test berdasarkan peluang kemunculan satu sisi koin dalam satu kali kejadian.
8. Pengujian hipotesisnya adalah dari tabel 14.5 di atas jumlah run untuk sebelah kanan adalah 8 dan sebelah kiri juga 8. Sehingga jumlah run keseluruhan berjumlah 16. Dengan jumlah sampel sebanyak 24 dengan

n1 adalah 12 dan n2 adalah 12 maka nilai terkecil (merujuk pada tabel nilai run test) adalah 7 sedangkan nilai terbesar adalah 19. Dari jumlah run sebanyak 16 terletak di antara 7 dan 19 makan kesimpulannya adalah Ho diterima. Dengan Ho diterima dan H1 ditolak makan wawancara bersifat random.

9. Kesimpulan akhir dari penelitian ini adalah siswa di SMA Jaya mempunyai variasi waktu dalam menentukan waktu terbaik untuk belajar. Waktu terbaik dipilih adalah setelah pulang sekolah dan malam hari. Sehingga, peluang memilih belajar pada waktu setelah pulang sekolah dan malam hari adalah sama yaitu 50 persen.

Contoh kasus apabila sampel dalam jumlah besar:

Penelitian dilakukan untuk mengetahui pria dan wanita dalam pemilihan kepala desa bersifat random atau tidak (artinya setiap kandidat kepala mempunyai peluang yang sama dipilih oleh pria dan wanita). Dari hasil pengamatan yang berhasil dikumpulkan dari antrian pertama hingga antrian terakhir diperoleh data sebagai berikut.

P	WW	PP	W	P	WW	PP	WW	P	W
P	WW	PP	WWW	P	W	P	W	P	W
PPP	W	PP	W	P	WWWW				

Berdasarkan data di atas tentukan kesimpulan yang dapat diberikan.

Jawaban:

1. Judul yang sesuai untuk penelitian adalah *Kecenderungan Pria dan Wanita dalam Memiliha Kepala Desa*
2. Peubah penelitian ini adalah pemilihan kepala desa
3. Jumlah sampel yang berhasil dikumpulkan adalah sebanyak (N) = 40 orang dengan kelompok pria (P) sebanyak 19 pria dan wanita (W) sebanyak 21 wanita.
4. Lokasi penelitian di kantor kepala desa

5. Hipotesis yang dapat diberikan dalam penelitian ini adalah:
   Ho: peluang setiap partai adalah sama 50 persen dipilih oleh pria dan 50 persen wanita dan urutan pemilihan bersifat random
   H1: peluang setiap partai adalah berbeda antar pria dan wanita dan urutan pemilihan bersifat tidak random atau berurutan
6. Syarat Pengujian Hipotesis adalah:
   Bila nilai p (dari tabel z hitung) lebih kecil atau sama dengan (≤) dari tingkat kesalahan yang ditetapkan (α) maka Ho diterima dan H1 ditolak.
7. Dari pengumpulan data pada soal makan jumlah run test yang diperoleh sebesar (r)=26.
8. Pengujian Hipotesis dapat dilakukan dengan mengaplikasikan nilai z seperti pada rumus. Dengan N sebanyak 40 responden dan tingkat kesalahan adalah 5% maka:

   $$z = \frac{r - \mu_r}{\sigma_r} = \frac{26 - \left(\frac{2\,.\,19\,.\,21}{19 + 21} + 1\right) - 0{,}5}{\sqrt{\frac{2\,.\,19\,.\,21\,(2\,.\,19\,.\,21 - 19 - 21}{(19 + 21)^2\,(19 + 21 - 1)}}}$$

   $$z = 1{,}78$$

   Setelah diperoleh nilai z adalah sebesar 1,78 makan nilai p berdasarkan pada tabel statistik adalah 0,0375 (angka diperoleh dari hasil uji dua pihak 1,7 dan 0,08 = 0,0375). Kesimpulannya adalah nilai p lebih kecil dari nilai harga α sebesar 5% (0,0375 < 0,05). Sehingga, nilai z yang lebih dari α maka Ho diterima dan H1 ditolak.
9. Kesimpulannya adalah urutan dalam antrian pemilihan kepala desa bersifat random serta peluang terpilihnya kandidat kepala desa adalah sama dari kelompok pria dan wanita.
10. Sarannya adalah kepada setiap kandidat kepala desa harus bisa lebih melakukan pendekatan kepada kedua kelompok dengan lebih seimbang.

14.3.4 Pengujian Hipotesis Interval/Rasio

1. Pengujian Rata-Rata

Prosedur uji hipotesis untuk rata-rata pada sampel besar adalah sebagai berikut:

1. Jenis Uji statistik adalah Uji Z
2. Hipotesis adalah:
   Ho: $\mu = \mu_0$ dan H1: $\mu \neq \mu_0$
   Ho: $\mu \leq \mu_0$ dan H1: $\mu > \mu_0$
   Ho: $\mu \geq \mu_0$ dan H1: $\mu < \mu_0$
3. Tingkat signifikansi atau α adalah 0.01, 0.05 dan 0.10.
4. Rumus untuk pengujian statistik menggunakan rumus z berikut:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

   Atau, apabila nilai tidak diketahui maka dapat menggunakan rumus berikut:

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

   Dengan syarat data berdistribusi normal dengan n yang besar.
5. Daerah kritis adalah:
   $Z < Z_{a/2}$ atau $Z > Z_{a/2}$ untuk H1: $\mu \neq \mu_0$
   $Z > Z_a$ untuk H1: $\mu > \mu_0$
   $Z < -Z_a$ untuk H1: $\mu < \mu_0$

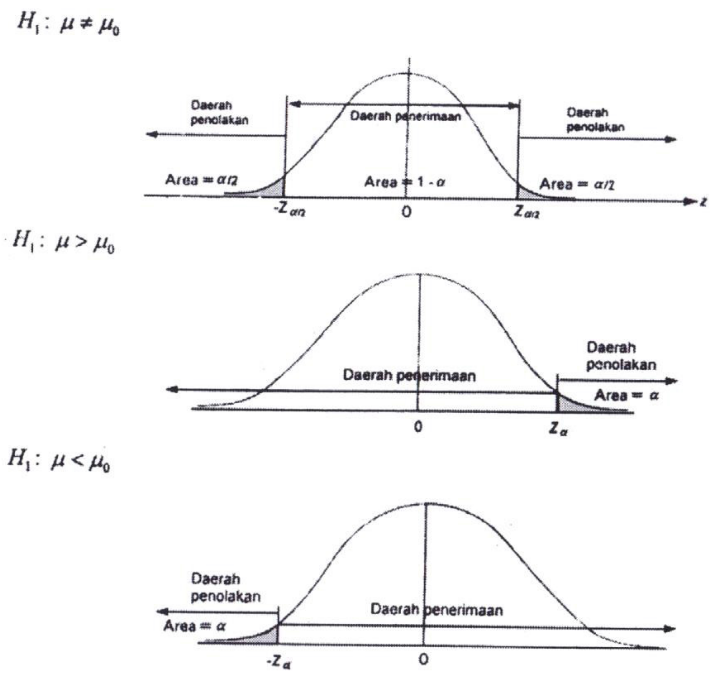
   Untuk nilai kritis dengan pengujian satu sisi (One-Tail) pada tingkat signifikansi 0.1, 0.05, dan 0.01 dapat dilihat pada Tabel 14.6 di bawah dengan andaian sampel adalah tak terhingga (∞).

TABEL 14.6 ***Nilai kritis z pada uji 1 arah***

α	H1: $\mu < \mu_0$	H1: $\mu > \mu_0$
0.01	$Z < -Z_{0.01}$ ($Z < -2.33$)	$Z > Z_{0.01}$ ($Z > 2.33$)
0.05	$Z < -Z_{0.05}$ ($Z < -1.64$)	$Z > Z_{0.05}$ ($Z > 1.64$)
0.10	$Z < -Z_{0.10}$ ($Z < -1.28$)	$Z < Z_{0.05}$ ($Z > 1.28$)

TABEL 14.7 ***Nilai kritis z pada uji 2 arah***

α	H1: $\mu < \mu_0$	H1: $\mu > \mu_0$
0.01	$Z < -Z_{0.005}$ ($Z < -2.58$)	$Z > Z_{0.01}$ ($Z > 2.58$)
0.05	$Z < -Z_{0.025}$ ($Z < -1.96$)	$Z > Z_{0.05}$ ($Z > 1.96$)
0.10	$Z < -Z_{0.05}$ ($Z < -1.64$)	$Z < Z_{0.05}$ ($Z > 1.64$)



GAMBAR 14.6 ***Daerah kritis pada uji hipotesis untuk satu nilai rata-rata***

Contoh:

Rata-rata waktu yang diperlukan untuk memproduksi kertas adalah 1.2 menit/rim. Telah dilakukan pengambilan sampel secara acak dengan 36 rim mempunyai rata-rata waktu produksi per rim nya adalah 1.5 menit. Jika diketahui nilai adalah 0.30 menit dengan α = 0.01. Berikan kesimpulan dari kasus ini berdasarkan pengujian hipotesis pada rata-rata waktu yang dibutuhkan untuk menghasilkan 1 rim kertas.

Jawaban:

1. Jenis uji statistik : uji *Z*.
2. Hipotesis : $H_{0:}$ $\mu \leq 1.2$ *versus* $H_{1:}$ $\mu > 1.2$
3. Tingkat signifikansi : $a = 0.01$
4. Daerah kritis : $Z > Z_{0.01}$, yaitu $Z > 2.33$
5. Statistik penguji:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

   yang berdistribusi normal standar.

   $\bar{x} = 1.5$ $\sigma = 0.30$ $n = 36$

   Standar error = $\sigma/\sqrt{n} = 0.30/\sqrt{36} = 0.05$

   Uji-z tersebut adalah:

$$z = \frac{1.5 - 1.2}{0.05} = 6.0$$

   yang terletak pada daerah kritis ($Zuji > 2.33$)

6. Kesimpulan: H_0 ditolak pada $\alpha = 0.01$

   Seandainya hendak dilakukan uji 2-sisi, perbedaannya adalah:

   a. Hipotesis:

      H**o**: $\mu = 1.2$ *vs* H1: $\mu \neq 1.2$

   b. Daerah kritis: $Z < -Z_{0.025}$, yaitu $Z < -2.58$

      atau $Z > Z_{0.025}$, yaitu $Z > 2.58$

   Dengan statistik penguji yang sama:

   $Z_{uji} = 6.0$,

yang juga terletak pada daerah kritis $Z > 2.58$, diperoleh kesimpulan yang sama, yaitu Ho ditolak pada $a = 0.01$.

2. Pengujian Proporsi

Prosedur uji hipotesis untuk rata-rata pada sampel besar adalah sebagai berikut:

1. Jenis Uji statistik adalah Uji Z
2. Hipotesis adalah:
   Ho: $p = p_0$ dan H1: $p \neq p_0$
   Ho: $p \leq p_0$ dan H1: $p > p_0$
   Ho: $p \geq p_0$ dan H1: $p < p_0$
3. Tingkat signifikansi atau α adalah 0.01, 0.05 dan 0.10.
4. Rumus untuk pengujian statistik menggunakan rumus z berikut:

$$z = \frac{p - p_0}{\sqrt{p_0 q_0 / n}}$$

   Dengan syarat data berdistribusi normal dengan n yang besar.
5. Daerah kritis adalah:
   $Z < Z_{a/2}$ atau $Z > Z_{a/2}$ untuk H1: $p \neq p_0$
   $Z > Z_a$ untuk H1: $p > p_0$
   $Z < -Z_a$ untuk H1: $p < p_0$

Contoh:
Diantara 900 petani sebagai sampel acak petani di DIY, 610 orang adalah buruh tani. Dengan a= 0.05akan diuji apakah proporsi buruh tani di DIY tidak kurang daripada 65%. Seandainya proporsi buruh tani pada populasi petani DIY melebihi 65%, diperlukan perubahan untuk memperbaiki dan meningkatkan taraf kehidupan populasi petani di DIY, maka uji hipotesis yang diperlukan adalah uji 1-sisi dengan hipotesis

$H_{0:}\ P \leq 0.65$ *vs* $H_{1:}\ P > 0.65$.
$n = 900$ $p = 0.678$
$p_o = 0.65$ $q_o = 1 - p_o = 1 - 0.65 = 0.35$
Standar error $= \sqrt{p_0 q_0 / n} = \sqrt{(0.65 \,.\, 0.35/900} = 0.016$

Pengujian Hipotesis (one-tail/1-sisi)

1. Jenis Uji statistik adalah Uji Z
2. Hipotesis adalah:

   Ho: $p \leq 0.65$ dan H1: $p > 0.65$
3. Tingkat signifikansi atau α adalah 0.05
4. Daerah kritis adalah: $Z>Z_{0.05}$ atau $Z>1.64$
5. Rumus untuk pengujian statistik menggunakan rumus z berikut:

$$z = \frac{p - p_0}{\sqrt{p_0 q_0 / n}}$$

   Dengan syarat data berdistribusi normal, dengan:

$$z = \frac{0.678 - 0.65}{0.016} = 1.75$$

   Terletak pada daerah kritis $Z_{0.05}>1.64$.
6. Kesimpulan dari hasil pengujian hipotesis ini adalah Ho ditolak dan H1 diterima yang artinya dengan tingkat signifikansi sebesar 5%, proporsi buruh tani di daerah DIY lebih besar daripada 65%.

Bab 15

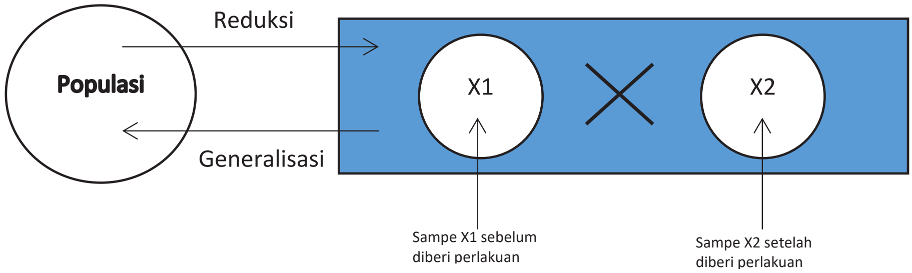
Pengujian Hipotesis Komparatif

15.1 PENGANTAR

Pada bab sebelumnya kita telah membahas apa itu pengujian hipotesis deskriptif. Di bab ini, pokok pembahasan akan menjelaskan tentang pengujian hiptesis komparatif. Pengujian hipotesis komparatif artinya melakukan perbandingan terhadap sampel-sampel berpasangan untuk melihat apakah terdapat perbedaan yang signifikan. Perlu dibatasi sampel yang dimaksud adalah dua sampel berpasangan dan dua sampel independen. Untuk mengetahui lebih lanjut hubungan antar nilai peubah maka dilakukan pengujian sesuai dengan jenis data yang diperoleh apakah sebaiknya menggunakan uji Mc Nemar pada data nominal, Sign test dan Wilcoxon pada data ordinal atau uji t-test pada dua sampel untuk data interval rasio.

15.2 HIPOTESIS KOMPARATIF

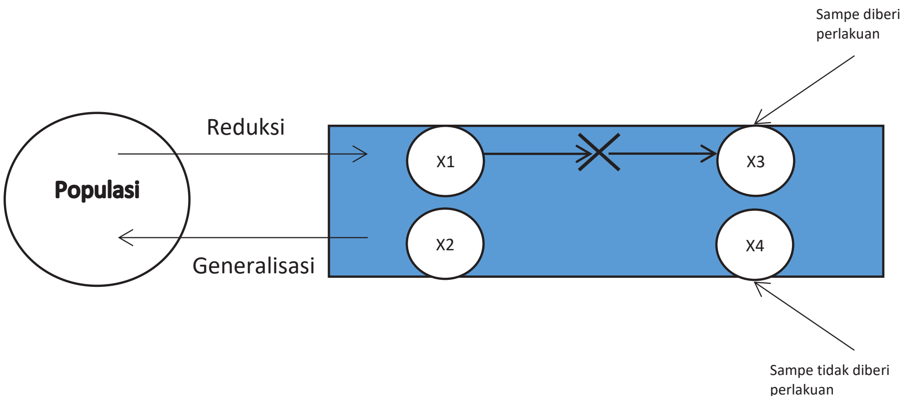
Dalam uji hipotesis komparatif pengujian populasi dilakukan dengan membandingkan nilai hitung dari sampel yang telah ditentukan. Sehingga, melalui uji ini kita dapat melihat nilai pengaruh signfikansi peubah independen terhadap peubah dependen. Sampel berpasangan tersebut dapat diilustrasikan seperti Gambar 15.1 berikut:



GAMBAR 15.1 ***Dua sampel berpasangan pada waktu sebelum dan setelah diberi perlakuan***

Dari gambar di atas contoh sederhana yang dapat diberikan ketika sampel yaitu masyarakat membeli produk barang setelah diberi iklan dan sebelum beri iklan. Hal ini berarti kedua sampel dilakukan dengan dua kali pengukuran. Setelah kedua kelompok masyarakat mendapat perlakuan berbeda dari iklan produk yang akan dibeli, maka kita dapat melihat sampel mana yang lebih banyak membeli untuk produk yang diiklankan. Anggota sampel X1 dicirikan sama dengan sampel pada X2. Sampel X1 pula adalah hasil observasi (pengamatan) sebelum produk diberi iklan, sedangkan sampel X2 merupakan hasil pengamatan setelah produk diberi iklan.

Contoh kasus lain untuk dua sampel berpasangan adalah sampel yang tidak diberi perlakuan dan diberi perlakuan. Jika pada kasus di atas sampel dinilai pada produk sebelum dan setelah diberi iklan. Pada bagian ini kelompok sampel akan diamati apakah cenderung membeli produk apabila produk yang dibeli diiklankan di media massa atau tidak diiklankan sama sekali. Sehingga, ilustrasi yang dapat diberikan seperti pada Gambar 15.2.



GAMBAR 15.2 ***Dua sampel berpasangan yang diberi perlakuan dan tidak diberi perlakuan***

Berdasarkan nilai hitung pengujian hipotesis komparatif ini, syarat yang terdapat dalam pengambilan keputusan adalah apabia nilai Ho diterima berarti kedua sampel yang dilihat dari nilai perbandingannya dapat digeneralisasikan sesuai dengan standar error yang telah ditetapkan. Hal ini berarti kedua sampel tidak mempunyai perbedaan atau hubungan terhadap perlakuan yang diberikan. Sebaliknya, jika H1 diterima berarti kedua sampel mempunyai perbedaan terhadap jumlah pembelian produk yang diberi iklan dengan produk yang tidak diiklankan.

Jenis pengujian ini tidak jauh berbeda dengan pengujian hipotesis deskriptif. Hal ini dapat dilihat dari pengambilan satu peubah penelitian tetapi meliputi sampel yang berbeda dalam populasi bahkan dilakukan pada waktu yang berbeda. Uji hipotesis komparatif juga menguji hubungan antara dua sampel atau lebih. Namun, pada bab ini fokus pembahasan akan lebih kepada pengujian dua sampel berpasangan dan dua sampel independen.

Pengujian hipotesis komparatif dua sampel berpasangan mempunyai beberapa jenis perhitungan statistik yang dapat diaplikasikan. Masing-masing perhitung bergantung pada jenis data dan kelompok sampelnya seperti yang ditunjukkan dalam tabel 15.1.

TABEL 15.1 ***Kategorisasi Perhitungan Pengujian Hipotesis Komparasi***

MACAM DATA	BENTUK KOMPARASI			
	DUA SAMPEL		K SAMPEL	
	Berpasangan	Independen	Berpasangan	Independen
Interval ratio	t-test *duasampel	t-test *duasampel	One Way Anova* Two Way Anova	One Way Anova* Two Way Anova
Nominal	Mc Nemar	Fisher Exact Chi Kuadrat Two sampel	Chi Kuadrat for k sampel Cochran Q	Chi Kuadrat for k sampel
Ordinal	Sign test Wilcoxon Matched Pairs	Median test Man-Whitney U test Kolomogorov Smirnov Wald-Wolfowitz	Friedman Two Way Anova	Median Extention Kruskal-Walls One Way Anova

Catatan: *Statistik parametrik

15.3 UJI HIPOTESIS KOMPARATIF DUA SAMPEL BERPASANGAN

15.3.1 Mc Nemar Test

Teknik ini digunakan untuk menguji hipotesis komparatif dua sampel yang berkorelasi bila datanya berbentuk niminal atau diskrit. Rancangan penilaian biasanya berbentuk " before after". Jadi hipotesis penelitian merupakan perbandingan antara nilai sebelum dan sesudah yang di dalamnya ada perlakuan. Sebagai panduan untuk menguji signifikansi setiap perubahan, maka data perlu disusun ke dalam table segi empat ABCD seperti berikut:

Sebelum	Sesudah	
	-	+
+	A	B
-	C	D

Tanda (+) dan (-) sekedar dipakai untuk menadai jawaban yang berbeda, jadi tidak harus bersifat positif dan negative. Kasus-kasus yang menunjukan perubahan antara jawaban pertama dan kedua muncul dalam sel A dan D. seseorang dicatat dalam cel A jika berubah dari tambah ke kurang, dan dicatat dalam cel D jika jika ia berubah dari kurang ke tambah. Jika tidak terjadi perubahan yang di observasi yang berbentuk tambah dia di catat di sel B, dan di catat di cel C bila tidak terjadi perubahan yang di observasi yang berbentuk kurang.

A + D adalah jumlah jumlah yang berubah, dan B dan C yang tidak berubah. Ho = ½ (A + D) berubah dalam suatu arah, dan merupakaxn frekuensi yang diharapkan di bawah Ho pada kedua buah sel yaitu A dan D. Test Mc Nemer berdistribusi Chi Kuadrat (x), oleh karena itu rumus yang digunakan untuk pengujian hipotesis adalah rumus chi kuadrat. Persamaan dasarnya ditunjukkan sebagai berikut:

$$X^2 = \sum_{i=1}^{K} \frac{(f_0 - f_h)^2}{f_h}$$

Dimana:

f_o = frekuensi yang diobservasi dalam kategori ke-i

f_h = frekuensi yang diharapkan dibawah fo dalam kategori ke-i

Setelah memperoleh nilai Chi Kuadrat tersebut, kita dapat melakukan uji signifikansi dengan menggunakan nilai peubah A dan D. Dimana A adalah jumlah observasi pada waktu A, sedangkan D adalah banyak kasusu yang diamati pada waktu D. Sehingga, rumus sederhana yang dapat digunakan seperti yang ditunjukkan pada rumus berikut:

$$X^2 = \frac{(A + D)^2}{A + D}$$

$$X^2 = \sum_{i=1}^{K} \frac{(f_0 - f_h)^2}{f_h} = \frac{(A + D)^2}{A + D} :$$

$$= \frac{\left(A - \frac{A + D}{2}\right)^2}{\frac{A + D}{2}} + \frac{\left(D - \frac{A + D}{2}\right)^2}{\frac{A + D}{2}}$$

Pengujian signifikansi semakin baik setelah diberi pengurangan angka 1 pada akhir rumus sesuai dengan adanya perbaikan "correction" dari ahli statistik Yates (1934). Pembetulan ini atau yang lebih dikenal dengan koreksi kontinuitas dilakukan karena distribusi data adalah normal. Sehingga, jenis data yang bersifat normal maka data tersebut adalah data kontinum. Adapaun rumus dari korelasi kontinuitas tersebut seperti berikut:

$$X^2 = \frac{(|A-D| - 1)^2}{A+D} \text{ dengan syarat } dk = 1$$

Contoh:

Perusahaan e-commerce produk kecantikan sedang melakukan penelitian terkait minta beli pelanggan terahadap barang yang telah diberi diskon pembelian. Dari populasi pengguna aplikasi tersebut, diambil secara acak sampel pelanggan yang melakukan pembelian produk sejumlah 200 pengguna. Sebelum diskon diberikan diketahui 150 pengguna tidak membeli produk dan 50 melakukan pembelian. Periode berikutnya, perusahaan memberikan diskon untuk produk yang sama dalam rentang waktu tertentu. Diperoleh data dari sampel yang sama, dari 200 pengguna sejumlah 125 orang membeli produk yang telah diberi diskon sedangkan 75 tidak melakukan pembelian. Dari angka 125 orang pembeli diketahi sebanyak 40 orang adalah pembeli tetap pada produk sebelum diberi diskon dan 85 orang yang tidak membeli sebelum produk didiskon. Kemudian dari 75 orang sisa yang tidak membeli terdiri 10 orang yang sebelum diskon membeli produk dan 65 tidak pernah melakukan pembelian produk baik sebelum maupun setelah diskon. Berdasarkan kasus di atas, apa keputusan terbaik yang harus diambil perusahaan tersebut.

Jawaban:

Berdasarkan informasi yang dipaparkan, kita dapat melakukan langkah-langkah penelitian seperti berikut:

1. Judul penelitian yang dapat diberikan untuk kasus di atas adalah pengaruh diskon harga terhadap penjualan produk kecantikan.
2. Dalam penelitian ini peubah yang digunakan adalah produk yang diberi diskon sebagai peubah independen dan jumlah peningkatan penjualan produk sebagai peubah dependen.

3. Adapun rumusan masalah penelitian ini adalah apakah terdapat pengaruh positif dan signifikan terhadap pemberian diskon harga terhadap penjualan produk kecantikan di perusahaan.
4. Dalam satu populasi penelitian ditetapkan jumlah sampel kajian sebanyak 200 orang pembeli yang sama sebelum dan setelah produk diberi diskon.
5. Hipotesis yang dapat diberikan adalah sebagai berikut:
   Ho: tidak adanya perbedaan jumlah penjualan produk sebelum dan sesudah diberi diskon
   H1: adanya perbedaan jumlah penjualan produk sebelum dan sesudah diberi diskon
6. kriteria pengujian hipotesis diberikan seperti berikut:
   Ho diterima apabila nila hitung Chi Kuadrat lebih kecil dari nilai tabel pada tingkat kesalahan (standard error) tertentu.
7. Tempat penelitian dilakukan melalui sistem user perusahaan terhadap pengguna aplikasi e-commerce terhadap proses pembelian yang berhasil dilakukan
8. Penyajian data dapat dilakukan dengan menyusun kembali informasi ke dalam tabel seperti berikut:

Penjualan Sebelum Diskon	Penjualan Setelah Diskon
50 orang pembeli	125 orang pembeli (40 tetap dan 85 pembeli baru)
150 orang tidak membeli	75 orang tidak membeli (10 berubah dan 65 tetap)
Jumlah 200 orang	Jumlah 200 orang sampel dengan kriteria 105 pembeli tetap dan 95 orang pembeli berubah

9. Pengujian hipotesis akan lebih mudah dilakukan dengan mengelompokkan kembali pembeli berdasarkan kriteria masing-masing serta menggunakan rumus perbaikan dari Yates.

Perilaku Pembeli	Penjualan Setelah Diskon	
Membeli	85 (A)	65 (B)
Tidak membeli	40 (C)	10 (D)

Menggunakan rumus perbaikan Yates, maka:

$$X^2 = \frac{(|A - D| - 1)^2}{A + D}$$

$$X^2 = \frac{(|85 - 10| - 1)^2}{85 + 10}$$

$$X^2 = 57{,}642$$

Nilai hitung Chi kuadrat di atas diperoleh sebesar 57,642. Kemudian, kita dapat melakukan perbandingan dengan menggunakan nilai tabel dengan tingkat kesahalan sebesar α= 5%. Sehingga, diperoleh nilai α= 3,894. Berdasarkan kedua nilai hitung chi kuadrat dan nilai tabel, maka kesimpulan sementara yang dapat ditarik adalah Ho diterima dan H1 ditolak. Hal ini karena nilai hitung chi kuadrat lebih besar berbanding nilai tabel (57,642 > 3,84).

10. Kesimpulan akhir yang dapat diberikan berdasarkan informasi yang diberikan serta hasil pengujian hipotesis yang telah dilakukan di atas maka didapati adanya perbedaan yang signifikan antara penjualan sebelum dan sesudah produk diberi diskon harga. Keputusan ini didasarkan pada penerimaan Ho dan penolak H1. Sehingga, diskon harga yang diberikan oleh perusahaan memberi pengaruh kepada pembeli mengakibatkan meningkatnya nilai penjualan produk. Oleh karena itu, perusahaan dipertimbangkan untuk melakukan diskon harga untuk periode berikutnya.

15.3.2 *Sign Test* (Uji Tanda)

Pengujian hipotesis berikutnya untuk jenis data ordinal adalah *sign test* atau lebih umunya disebut uji tanda. Penamaan uji tanda ini diberikan mempunyai maksud bahwa semua data yang dianalisis dinyatakan dalam bentuk tanda positif (+) atau negatif (-). Sebagai contoh dalam hal percobaan penerapan jalur satu arah bagi pengemudi mobil dan sepeda motor. Selama percobaan dilakukan hasil tidak dinyatakan dalam bentuk angka atau kuantitatif melainkan dinyatakan apakah percobaan memberi dampak positif atau negatif terhadap pengurangan kemacetan di jalan. Contoh lain adalah pemberian bonus kepada pegawai. Apakah setelah bonus diberikan kepada pegawai perusahaan memberi pengaruh positif terhadap efektivitas dan menguntungkan bagi perusahaan. Sehingga dalam proses penelitiannya tidak membahas tentang besaran pengaruh secara angka tetapi lebih kepada pernyataan berpengaruh positif atau negatif.

Fungsi uji tanda dalam percobaan bertujuan melihat pengaruh suatu peubah atau perubahan selama proses eksperimen berlangsung pada keadaan tertentu. Keadaan atau kondisi tersebut menurut John W. Best, adalah: Jika penilaian atas efek peubah atau perlakuan eksperimen tidak dapat diukur, tetapi hanya dapat dinilai dengan sistem juri dalam bentuk performansi baik atau jelek, superior atau inteferior dsb. Jika anggota-anggota kelompok eksperimen dan kelompok kontrol terdiri dari 10 pasangan atau lebih, yang di pasangkan atas dasar IQ; bakat, saudara kembar atau dasar-dasar pemasangan lainya. Subjek bisa jadi dipasangkan dengan sendirinya, menurut pola pre-observasi dan post-obserasi. Pada`suatu ketika, mereka bertindak sebagai kelompok kontrol (yakni pada saat per-observasi), dan pada sat yang lain menjadi kelompok eksperimen (yakni pada saat eksperimen).

Perlu diingat dalam penggunaan rumus *sign test* adalah bahwa teknik ini sangat tepat digunakan untuk menganalisa perbedaan antara sample-sample terikat, bukan sample bebas (dependen). Disamping itu perlu juga dipahami bahwa tes ini tidak menunjukan besarnya perbedaan, akan tetapi hanya menilai arah superior atau interior. Pada sampel kecil berjumlah ≤ 25 maka pengujian dilakukan dengan menggunakan prinsip distribusi Binomial dengan $P=Q=0.5$ di mana N = banyak pasangan. Apabila pada suatu pasangan dalam observasi tidak mempunyai perbedaan yaitu selisihnya = 0, maka pasangan tersebut tidak dapat digunakan dalam proses analisis. Hal ini menyebabkan pengurangan dari jumlah N awal yang telah ditetapkan.

Contoh:

Perusahaan jasa keuangan sedang melakukan observasi terhadap kesejahteraan keluarga pegawainya. Dengan kenaikan gaji yang diberikan apakah memberi pengaruh positif atau negatif pada tingkat kesejahteraannya. Secara acak dipilih 20 pasangan suami dan istri. Masing-masing dari mereka diminta untuk mengisi kuesioner. Dari hasil pengumpulan data diperoleh hasil seperti yang ditunjukkan dalam tabel berikut:

Jawaban Suami (Kenaikan Gaji)		Jawaban Istri (Kenaikan Gaji)	
Sebelum	Setelah	Sebelum	Setelah
2	4	2	3
1	3	1	2
2	6	3	8
4	5	2	4
1	6	3	7
5	9	6	9
2	6	3	5
2	7	2	4
6	9	5	9
4	8	4	7
2	3	7	9
1	4	1	6
2	7	2	6
1	4	1	3
3	6	2	4
2	4	4	5
6	7	5	7
2	3	4	6
4	6	2	3
1	6	2	4

Jawaban:

1. Judul penelitian yang tepat untuk penelitian di atas adalah Pengaruh kenaikan gaji pegawai terhadap kesejahteraan atau adakah perbedaan sebelum dan setelah kenaikan gaji pegawai.
2. Peubah yang digunakan adalah kenaikan gaji sebagai peubah independen dan kesejahteraan sebagai peubah dependen
3. Rumusan masalah penelitian ini adalah adakah pengaruh kenaikan gaji terhadap kesejahteraan pegawai?
4. Ukuran sampel terdiri dari 20 pasang suami dan istri yang dipilih secara acak.

5. Tempat penelitian dilakukan di kantor perusahaan dengan mengundang setiap pasangan dari para pegawai untuk datang mengisi kuesioner yang diberikan.
6. Hipotesis penelitian ini adalah sebagai berikut:
   Ho: Tidak adanya perbedaan yang signifikan pada kenaikan gaji pegawai terhadap kesejahteraan keluarga baik menurut suami atau pun istri
   H1: Tidak adanya perbedaan yang signifikan pada kenaikan gaji pegawai terhadap kesejahteraan keluarga baik menurut suami atau pun istri
7. Pengujian hipotesis harus memenuhi kriteria utama di mana Ho diterima apabila nilai hitung Chi kuadrat lebih kecil dari Chi kuadrat tabel
8. Sebelum dilakukan perhitungan langkah pertama yang dapat dilakukan adalah dengan menyaji data ke dalam tabel seperti berikut:

Jawaban Suami (Kenaikan Gaji)				Jawaban Istri (Kenaikan Gaji)			
Sebelum	**Setelah**	**Beda**	**Rank**	**Sebelum**	**Setelah**	**Beda**	**Rank**
2	4	2	4	2	3	1	5
1	3	2	4	1	2	1	5
2	6	4	2	3	8	5	1
4	5	1	5	2	4	2	4
1	6	5	1	3	7	4	2
5	9	4	2	6	9	3	3
2	6	4	2	3	5	2	4
2	7	5	1	2	4	2	4
6	9	3	3	5	9	4	2
4	8	4	2	4	7	3	3
2	3	1	5	7	9	2	4
1	4	3	3	1	6	5	1
2	7	5	1	2	6	4	2
1	4	3	3	1	3	2	4
3	6	3	3	2	4	2	4
2	4	2	4	4	5	1	5
6	7	1	5	5	7	2	4
2	3	1	5	4	6	2	4
4	6	2	4	2	3	1	5
1	6	5	1	2	4	2	4

Berdasarkan pengumpulan data di atas kita dapat melihat perubahan melalui perubahan pada kolom beda dan rank. Perhitungan beda dan rank ditentukan dengan menggunakan selisih dari nilai sebelum dan sesudah. Misalnya pada responden 1 dari pihak suami untuk sebelum kenaikan gaji memberikan penilaian 2 kemudian setelah kenaikan penilaian naik menjadi 4. Perbedaan ini menghasilkan angka 2, maka angka 2 ini dari urutan peringkat berada pada peringkat ke 4 (Angka 1 2 3 4 5 dalam peringkat 5 4 3 2 1). Dari data istri sebelum kenaikan memberi angka 2 dan setelah kenaikan memberi angka 3, selisih dari kedua angka adalah 1 yang berarti peringkat dalam rangking adalah 5. Hal ini berarti perubahan yang terjadi sangat kecil.

9. Pengujian hipotesis dapat dilakukan menggunkan Uji Tanda dengan menganalisis data ordinal atau berbentuk peringkat tersebut.

No.	Tingkat Perubahan		Arah			Tanda
	Suami	Istri				
1	4	5	4	<	5	+
2	4	5	4	<	5	+
3	2	1	2	>	1	-
4	5	4	5	>	4	-
5	1	2	1	<	2	+
6	2	3	2	<	3	+
7	2	4	2	<	4	+
8	1	4	1	<	4	+
9	3	2	3	>	2	-
10	2	3	2	<	3	+
11	5	4	5	>	4	-
12	3	1	3	>	1	-
13	1	2	1	<	2	+
14	3	4	3	<	4	+
15	3	4	3	<	4	+
16	4	5	4	<	5	+
17	5	4	5	>	4	-
18	5	4	5	>	4	-

No.	Tingkat Perubahan		Arah			Tanda
	Suami	Istri				
19	4	5	4	<	5	+
20	1	4	1	<	4	+

Berdasarkan tabel di atas didapati tanda positif dan negatif pada tabel sebanyak 13 tanda positif (+) dan 7 tanda negatif (-). Menggunakan tabel binomial dengan N=20 serta dengan nilai harga p paling kecil adalah 7, maka nilai p tabel (sampel terkecil) = 0,132. Dengan tingkat kesalahan (standard error) 5% maka kesimpulan sementara adalah nilai hitung 0,132 lebih besar berbanding 0,05. Sehingga, dari nilai perbandingan ini Ho diterima dan H1 ditolak. Kesimpulan yang dapat diberikan dari perhitungan pengujian hipotesis ini adalah tidak terdapat pengaruh yang positif dan signifikan dari naiknya gaji pegawai terhadap kesejahteraan keluarga baik menurut suami ataupun istri.

10. Kesimpulan akhir untuk penelitian pada kasus ini adalah tidak adanya pengaruh yang positif dan signifikan kenaikan gaji terhadap kesejahteraan keluarga pegawai perusahaan. Kalaupun dalam data terdapat sampel yang berpengaruh positif, pengaruh hanya terjadi pada sampel tersebut dan tidak dapat digeneralisasi ke dalam populasi dari semua pegawai perusahaan.
11. Saran dan rekomendasi untuk penelitian ini adalah sebaiknya dilakukan kajian lebih mendalam sebab kenaikan gaji tidak memberi pengaruh positif terhadap kesejahteraan pegawai. Kemungkinan nilai dari kenaikan gaji adalah kecil sehingga tidak dapat meningkatkan kesejahteraan.
12. Untuk memastikan hasil pengujian menggunakan Uji tanda apakah sudah tepat atau belum, maka dapat dilakukan pengujian hipotesis menggunakan rumus chi kuadrat dengan rumus seperti berikut:

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

Dimana:

n_1 = banyaknya data positif

n_2 = banyaknya data negatif

Sehingga, dari rumus ini kita dapati:

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

$$X^2 = \frac{[(13 - 7) - 1]^2}{13 + 7}$$

$$X^2 = \frac{25}{20} = 1.25$$

Untuk membuktikan apakah nilai Ho ditolak dan H1 diterima maka kita menggunakan nilai Chi kuadrat dengan dk = 1. Dengan standard error sebesar 5% atau 0,05 dan nilai tabel chi kuadrat sebesar 3.481 maka dapat disimpulkan bahwa nilai hitung chi kuadrat 1.25 < dari nilai tabel chi kuadrat 3.481. (1.25 < 3.481). Dengan demikian Ho diterima dan H1 ditolak. Sehingga hasil yang diperoleh dengan rumus chi kuadrat adalah sama dengan pengujian Uji Tanda pertama, yaitu kenaikan gaji tidak memberi pengaruh terhadap peningkatan kesejahteraan pegawai perusahaan.

15.3.3 Wilcoxon Match Pairs Test

Pengujian hipotesis komparatif menggunakan uji wilcoxon merupakan penyempurnaan dari uji tanda (*sign test*). Hal ini karena pada uji wilcoxon besarnya selisih angka pada tanda sampel positif atau negatif harus diperhitungkan, sedangkan pada uji tanda ini tidak perlu dilakukan. Pengembangan uji wilcoxon dari uji tanda dilakukan karena tingkat ketelitian uji ini lebih baik terutama untuk menunjukkan adanya perbedaan dari masing-masing sampel yang dibandingkan. Namun, sama halnya dengan uji tanda, uji wilcoxon dapat digunakan untuk menguji jenis data ordinal dengan pengujian hipotesis komparatif dua sampel yang saling berhubungan. Dalam rumus wilcoxon terdapat rumus *T* yang melambangkan jumlah peringkat dari sampel bernotasi positif (+) atau negatif (-) yang minoritas. Secara umum rumus yang selalu digunakan adalah seperti berikut:

$$z = \frac{T - \mu_T}{\sigma_T}$$

Contoh:
Di ruangan belajar mahasiswa Fakultas Ekonomi dan Bisnis kampus ciputat sedang melakukan penelitian tentang pengaruh AC terhadap fokus belajar mahasiswa. Pengumpulan data dilakukan sebanyak dua kali, sebelum dan sesudah AC dipasang. Fokus belajar mahasiswa sebelum dipasang AC dinotasikan dengan X1 dan sesudah dipasang AC dilambangkan dengan X2. Secara acak mahasiswa yang diambil sebagai sampel sebanyak 10 mahasiswa. Tentukan kesimpulan dan saran yang dapat diberikan.

No.	X1 (Sebelum)	X2 (Sesudah)
1	89	85
2	98	94
3	76	78
4	90	98
5	87	90
6	89	85
7	77	86
8	92	87
9	78	80
10	82	83

Jawaban:

1. Judul penelitian untuk penelitian di atas adalah pengaruh AC terhadap fokus belajar mahsiswa
2. Peubah yang digunakan adalah Pemasangan AC sebelum dan sesudah sebagai peubah independen dan fokus mahasiswa sebagai peubah dependen
3. Rumusan masalah penelitian adalah apakah pengaruh AC dapat meningkatkan fokus belajar mahasiswa di fakultas ekonomi dan bisnis?
4. Jumlah sampel dalam penelitian adalah sebanyak 10 orang
5. Penelitian dilakukan melalui metode pengisian kuesioner dan berdasarkan capaian nilai mahasiswa di akhir semester selama 1 semester. Analisis dilakukan dengan membanding nilai UTS dan UAS yang diperoleh mahasiswa dengan bobot penilaian yang sama.

6. Hipotesis penelitian ini adalah:
   Ho: Tidak adanya perbedaan fokus mahasiswa sebelum dan sesudah AC dipasang. Jadi AC tidak mempengaruhi fokus mahasiswa selama belajar di kelas
   H1: Adanya perbedaan fokus mahasiswa sebelum dan sesudah AC dipasang. Jadi AC tidak mempengaruhi fokus mahasiswa selama belajar di kelas
7. Kriteria pengujian hipotesis dapat dilakukan dengan ketentuan Ho diterima apabila nilai jumlah jenajng terkecil *T* (hitung) lebih besar dari nilai *T* tabel (pada tabel wilcoxon).
8. Analisis dapat dilakukan dengan terlebih dahulu menyajikan data seperti berikut dan kemudian menghitung nilai beda dan tanda jenjang dari masing-masing sampel.

No.	**X1 (Sebelum)**	**X2 (Sesudah)**	**Beda**	**Tanda Jenjang**		
				Jenjang	**+**	**–**
1	89	85	(–) 4	6	0	6
2	98	94	(–) 4	6	0	6
3	76	78	(+) 2	2.5	2.5	0
4	90	98	(+) 8	9	9	0
5	87	90	(+) 3	4	4	0
6	89	85	(–) 4	6	0	6
7	77	86	(+) 9	10	10	0
8	92	87	(–) 5	8	0	8
9	78	80	(+) 2	2.5	2.5	0
10	82	83	(+) 1	1	1	0
Jumlah					*T* = 29	26

9. Setelah memperoleh hasil nilai T seperti yang ditunjukkan pada tabel di atas, maka langkah selanjutnya adalah melakukan perbandingan dengan menggunakan nilai pada tabel *T* dengan sampel n=10 pada tingkat kesalahan 5% (uji dua pihak). Diperoleh nilai tabel *T* = 8. Sehingga, dengan nilai hitung *T*=26, maka Ho diterima dan H1 ditolak (26>8).

10. Untuk melihat keakuratan keputusan, langkah berikutnya dalam menggunakan rumus berikut:

$$z = \frac{T - \mu_T}{\sigma_T}$$

Karena sampel di bawah dari 25, maka diasumsikan distribusi data adalah normal. Kemudian kita dapat menghitung nilai *T* seperti berikut:

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

$$\mu_T = \frac{10(10+1)}{4} = 27.5$$

Selanjutnya, kita dapat menghitung nilai menggunakan rumus berikut:

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

$$\sigma_T = \sqrt{\frac{10(10+1)(2\,.\,10+1)}{24}} = 9.8$$

Maka, substitusikan masing-masing rumus ke dalam rumus *z* berikut:

$$z = \frac{T - \mu_T}{\sigma_T}$$

$$z = \frac{26 - 27.5}{9.8}$$

$$z = -\,0.153$$

Dengan nilai hitung *z* adalah sebesar 0.153 dengan tigkat kesalahan (standard error) sebesar 5% (pada dua pihak) maka diperoleh nilai *z* tabel = 0.4404. Dengan demikian Ho diterima dan H1 ditolak 0.153 < 0.4404). Sehingga, AC tidak memberikan pengaruh signifikan terhadap peningkatan fokus belajar mahasiswa di kelas.

11. Kesimpulan dan Saran
    Berdasarkan dari hasil kedua perhitungan di atas, baik uji wilcoxon maupun nilai uji *z test*, maka diperoleh kesimpulan akhir bahwa kelas yang diberi AC tidak memberikan pengaruh terhadap peningkatan fokus mahasiswa selama proses belajar mengajar. Oleh karena itu, saran dari penelitian ini adalah sebaiknya tidak perlu dipasang AC diruangan kelas mahasiswa Fakultas Ekonomi dan Bisnis.

15.4 UJI HIPOTESIS KOMPARATIF DUA SAMPEL INDEPENDEN

Pengujian hipotesis dua sampel independen berati menguji signifikansi antara dua sample yang tidak berpasangan dengan melihat perbedaan nilai keduanya. Biasanya penelitian yang selalu menggunakan uji hipotesis ini adalah penelitian berbasis data survei. Statistik yang digunakan untuk pengujian hipotesis komparatif ini adalah Chi Kuadrat, Fisher Exact Probability Test untuk data berjenis nominal dan Median Test, Mann Whitney-U Test, Kolmogorov-Smirnov, dan Wolfowits untuk jenis data ordinal.

15.4.1 Chi Kuadrat (X^2) Dua Sampel

Chi kuadrat digunakan untuk menguji hipotesis kompatatif dua sampel bila datanya berbentuk nominal dan sampelnya besar. Cara perhitungan dapat menggunakan rumus yang telah ada, atau dapat menggunakan Tabel Kontingensi 2 x 2 (dua baris x dua kolom). pada bab ini diberikan contoh penggunaan tabel kontingensi untuk menghitung harga chi kuadrat, karena lebih mudah.

Tabel Kontingensi

Sampel	Frekuensi Pada		Jumlah sampel
	Obyek I	Obyek II	
Sampel A	A	B	a+b
Sampel B	C	D	c+d
Jumlah	a+c	b+d	N

n = jumlah sampel

dengan memperhatikan koreksi yates, rumus yang digunakan untuk menguji hipotesis, adalah sebagai berikut (rumus 4.1)

$$x^2 = \frac{n\left(|ad - bc| - \frac{1}{2}n\right)^2}{(a+b)(a+c)(b+d)(c+d)}$$

Contoh 1:

Tim peneliti ingin mengetahui peluang calon bupati di kabupaten X. Setelah melalui proses pendaftaran dan penetapan calon, ditetapkan calon yang maju dalam kontes pemiliha kepala daerah adalah calon A dan B. Setelah survei dilakukan yang melibatkan 150 orang pemilih yang diambil secara acak dari 5 kelurahan berbeda. Dari hasil pengumpulan data yang mana setiap sampel dibagi sama rata, diperoleh angka 60 orang memilih calon A dan 50 orang memilih untuk calon B. Sedangkan yang tidak memilih calon A adalah 15 orang dan 25 orang tidak memilih calon B. Data data tersedia berikan kesimpulan dan saran dari hasil analisis yang dilakukan.

Jawaban:

1. Judul penelitian dapat dirumuskan adalah peluang terpilihnya calon bupati di kabupaten X
2. Peubah penelitiannya adalah calon bupati
3. Rumusan masalah

   Apakah terdapat perbedaan peluang antara calon bupati A dan B terpilih?
4. Sampel penelitian terdiri dari dua kelompok masyarakat yang setuju dan tidak. Kedua calon mendapatkan jumlah sampel yang sama yaitu 75 orang.
5. Hipotesis yang diberikan adalah:

   Ho: peluang calon bupati A dan B terpilih adalah sama sehingga tidak ada perbedaan antara kedua calon di kalangan masyarakat.

   H1: peluang calon bupati A dan B terpilih adalah berbeda sehingga terdapat perbedaan antara kedua calon di kalangan masyarakat.
6. Kriteria pengujian hipotesis penelitian adalah Ho diterima apabila nilai hitung Chi kuadrat lebih kecil kecil dari nilai tabel Chi kuadrat, dengan dk = 1 dan taraf kesalahan tertentu.
7. Sebelum dilakukannya analisis alangkah baiknya menyusun kembali hasil survei ke dalam tabel seperti berikut:

Frekuensi Masyarakat Memilih Calon A dan B

Sampel	Jenis Bank		Jumlah Sampel
	Memilih	Tidak Memilih	
Calon A	60	15	75
Calon A	50	25	75
Jumlah	110	40	150

8. Berdasarkan tabel pada poin 7 di atas, maka kita dapat melakukan perhitungan untuk memperoleh nilai hitung Chi Kuadrat kedua calon.

$$X^2 = \frac{n\left(|ad - bc| - \frac{1}{2}n\right)^2}{(a+b)(a+c)(b+d)(c+d)}$$

$$X^2 = \frac{150\left(|60\,.\,25 - 15\,.\,50| - \frac{1}{2}.\,150\right)^2}{(60+15)(60+50)(15+25)(50+25)}$$

$$X^2 = \frac{150\left(|60\,.\;25 - 15\,.\;50| - \frac{1}{2}.\,150\right)^2}{(60+15)(60+50)(15+25)(50+25)}$$

$$X^2 = 2.761$$

Pada tingkat kesalahan yang ditetapkan sebesar 5% dengan dk = 1, maka diperoleh nilai X^2 tabel = 3.481. Kesimpulan sementara adalah berdasarkan nilai X^2 = 2.761 dengan nilai tabel 3.481, maka Ho diterima dan H1 ditolak. Hal ini karena nilai X^2 masih lebih kecil dari nilai tabel X^2 yaitu 2.761 < 3.481.

9. Kesimpulan akhir yang dapat diberikan dari hasil penelitian di atas adalah terdapat perbedaan di kalangan masyarakat dalam memilih calon bupatinya. Hal ini berarti peluang kedua calon dipilih oleh masyarakat adalah sama dan mempunyai kelompok pemilih yang sama.
10. Saran dari penelitian untuk kedua calon bupati A dan B adalah sebaiknya lebih meningkatkan kampanye dan mencari dukungan tambahan dari kelompok masyrakat dari kelurahan berbeda dari sampel penelitian yang telah digunakan sebagai sumber data.

15.4.2 Fisher Exact Probability Test

Selain uji Chi Kuadrat, pengujian hipotesis pada dua sampel independen juga dapat menggunakan uji Fisher Exact Probability Test atau yang biasa dikenal dengan efek fisher. Kegunaan uji tes ini untuk melihat signifikansi peubah pada hipotesis sampel kecil atau kurang dari 20 dengan syarat data yang diteliti merupakan data nominal. Supaya memudahkan pemahaman dalam pengujiannya, maka hasil pengamatan sementara sebaiknya disusun dan ditabulasikan dalam tabel kontingensi (2x2) seperti berikut:

Tabel Kontingensi

Sampel	Frekuensi		Jumlah Sampel
	Objek X1	Objek X2	
1	A	B	A + B
2	C	D	C + D
Jumlah	A + C	B + D	N

Frekuensi menendakan bahwa objek yang diteliti mempunyai klasifikasi A dan B, misal seperti Setuju atau Tidak Setuju, Laki-Laki atau Perempuan dan lain-lain. Nilai data pada objek penelitian A B C dan D harus data nominal yang mempunyai nila frekuensi. Hal ini bertujuan untuk memudahkan perhitungan dalam rumus berikut:

$$\rho = \frac{(A+B)!\,(C+D)!!\,(A+C)!\,(B+D)!}{N!\,A!\,B!\,C!\,D!}$$

Dalam pengujian hipotesisnya kriteria yang harus dipenuhi adalah apabila nilai hitung $\rho > \rho$ tabel pada nilai signifikansi yang ditetapkan, maka Ho diterima dan H1 ditolak.

Contoh:

Perusahaan otomotif ingin melihat kecenderungan pembeli dalam memilih mobil berdasarkan warnanya. Asumsi yang beredar bahwa pembeli laki-laki lebih memilih mobil berwarna gelap sedangkan perempuan berwarna cerah. Untuk melihat apakah asumsi ini benar, maka dilakukan pengumpulan data sebagai sampel penelitian yang dilakukan secara acak. Dari 8 laki-laki yang membeli mobil diketahui 5 membeli mobil berwarna gelap dan 3 berwarna cerah. Kelompok perempuan dari 7 orang yang diamati, diperoleh 2 orang membeli warna gelap dan 5 mobil berwarna cerah. Tentukan kesimpulan dan saran yang dapat diberikan dari hasil analisis kasus tersebut.

Jawaban:

1. Judul penelitian untuk penelitian di atas adalah kecenderungan pembeli terhadap pembelian mobil berdasarkan warna
2. Peubah penelitiannya adalah warna mobil
3. Rumusan masalah penelitian adalah apakah terdapat perbedaan antara laki-laki dan perempuan dalam membeli mobil berdasarkan warnanya.
4. Jumlah sampel yang berhasil dikumpul berjumlah 8 laki-laki dan 7 perempuan
5. Hipotesis yang dapat diberikan adalah seperti berikut:
   Ho: Tidak adanya perbedaan baik laki-laki maupun perempuan dalam membeli warna gelap atau cerah
   H1: Adanya perbedaan baik laki-laki maupun perempuan dalam membeli warna gelap atau cerah
6. Dalam proses pengujian hipotesis maka kriteria yang harus dipenuhi adalah Ho diterima apabila nilai hitung dari ρ > dari nilai hitung ρ tabel pada tingkat kesalahan tertentu.
7. Langkah berikutnya untuk memudahkan perhitungan maka tabulasikan informasi di atas ke dalam tabel seperti berikut:

Pembelian Mobil Berdasarkan Warna

Sampel	Frekuensi		Jumlah Sampel
	Gelap	Cerah	
Laki-Laki	5	3	8
Perempuan	2	5	7
Jumlah	7	8	15

Setelah memasukkan semua informasi ke dalam tabel di atas, maka perhitungan dapat dilakukan dengan menggunakan rumus seperti berikut:

$$\rho = \frac{(5+3)!\,(2+5)!\,(5+2)!\,(3+5)!}{15!\,5!\,3!\,2!\,5!}$$

$$\rho = \frac{40320.5040.5040.40320}{1307674368000.120.6.2.120} = 0{,}82$$

Dengan tingkat kesalahan sebesar 5%, maka nilai ρ pada tabel diperoleh sebesar 0.37. Hasil perbandingan dari kedua nilai ρ adalah Ho diterima dan H1 ditolak. Ini disebabkan oleh nilai hitung ρ lebih besar dari nilai ρ tabel (0.37 > 0.05). Keputusan dari hasil pengujian hipotesis adalah tidak terdapat perbedaan antara pembeli laki-laki dan perempuan dalam membeli mobil berdasarkan warnanya.

8. Kesimpulan dari penelitian ini adalah baik kelompok pembeli laki-laki dan kelompok pembeli perempuan tidak terlalu mempertimbangkan warna ketika membeli mobil baik mobil dengan warna gelap atau cerah.
9. Saran kepada perusahaan berdasarkan hasil penelitian ini adalah sekiranya perusahaan otomotif berniat ingin melakukan promosi dan iklan mobil maka warna mobil bukan faktor penting dalam keputusan pembelian.

15.4.3 Test Median (*Median Test*)

Test Median digunakan untuk menguji signifikansi hipotesis komparatif dua sampel independen bila datanya berbentuk nominal atau ordinal. Pengujian didasarkan atas median dari sampel yang diambil secara random. Dengan demikian Ho yang akan diuji berbunyi. Tidak terdapat perbedaan dua kelompok populasi berdasarkan mediannya. Kalau Test Fisher digunakan untuk sampel kecil dan Test Chi Kuadrat digunakan untuk sampel besar, maka test median ini digunakan untuk sampel antara Fisher dan Chi Kuadrat. Berikut ini diberikan panduannya.

- Jika $n_1 + n_2 > 40$, dapat dipakai tes Chi Kuadrat dengan koreksi kontinuitas dari Yates.

- Jika $n_1 + n_2$ antara 20 – 40 dan jika tak satu sel pun memiliki frekuensi yang diharapkan 5, dapat digunakan Chi Kuadrat dengan koreksi kontinuitas. Bila f < 5 maka dipakai tes Fisher.
- Kalau $n_1 + n_2 > 20$ maka digunakan tes Fisher.

  Untuk menggunakan tes media maka pertama-tama harus dihitung gabungan dua kelompok,(median untuk semua kelompok). Selanjutnya dibagi dua. Dan dimasukkan ke dalam tabel seperti berikut:

Kelompok	Kel I	Kel II	Jumlah
Diatas median gabungan	A	B	A+B
Dibawah median gabungan	C	D	C-D
Jumlah	A+C = n_1	B+D = n_2	N = n_3

Dimana:

A = Banyak kasus dalam kelompok I di atas median gabung = $\frac{1}{2}n_1$

B = Banyak kasus dalam kelompok II di atas median gabungan = $\frac{1}{2}n_2$

C = Banyak kasus dalam kelompok I di bawah median = $\frac{1}{2}n_1$

D = Banyak kasus dalam kelompok II di bawah median = $\frac{1}{2}n_2$

Pengujian dapat menggunakan rumus Chi Kuadrat seperti ditunjukkan pada rumus 4.4 berikut.

$$x^2 = \frac{N\left[(AD - BC) - \frac{N}{2}\right]}{(A+B)(C+D)(A+C)(B+D)}$$

Rumus diatas derajat kebebasan (dk) = 1

Contoh:

Dilakukan penelitian untuk mengetahui apakah penghasilan para nelayan berbeda dengan para petani berdasarkan mediannya. Berdasarkan wawancara terhadap para 10 petani dan 9 nelayan diperoleh data tercantum pada tabel 4.5 berikut.

TABEL 4.5 *Penghasilan petani dan nelayan (X 1000 rupiah)*

No	Petani	Nelayan
1	50	45
2	60	50
3	70	55
4	70	60
5	75	65
6	80	65
7	90	70
8	95	80
9	95	100
10	100	

Berdasarkan hal tersebut diatas maka

1. Judul penelitiannya dapat dirumuskan sebagai berikut:
   Perbedaan penghasilan kelompok petani dan nelayan
2. Peubah penelitiannya adalah: penghasilan
3. Rumusan Masalah
   Adakah perbedaan yang signifikan antara penghasilan kelompok petani dan nelayan?
4. Sampel
   Dua kelompok sampel yaitu petani dengan jumlah 10 orang dan nelayan dengan jumlah 9 orang.
5. Hipotesis:
   H_0: tidak terdapat perbedaan yang signifikan antara penghasilan petani dan nelayan.
   H_1: terdapat perbedaan yang signifikan antara penghasilan petani dan nelayan.
6. Criteria Pengujian Hipotesis
   H_0: diterima bila Chi Kuadrat hitung < tabel.
   H_0: diterima bila Chi kuadrat hitung ≥ table.
7. Penyajian Data
   Untuk menghitung median gabungan maka data dua kelompok tersebut disusun dari yang kecil menuju yang besar, yaitu sebagai berikut:

   45 50 50 55 60 60 65 65 70 70 70 75 80 80 90 95 95 100 100

Median (nilai tengah) untuk kelompok tersebut jatuh pada urutan ke 10, yang nilainya = 70.

Dari table diketahui bahwa A = 6, C = 4, B = 2 dan D = 7.

Harga-harga tersebut kemudian dimasukkan ke dalam table.

Jumlah skor	Petani	Nelayan	Jumlah
Di atas Median Gabungan	A = 6	B = 2	A + B = 8
Di bawah Median Gabungan	C = 4	D = 7	C + D = 11
Jumlah	10	9	N = 19

8. Perhitungan

   Harga-harga dalam tabel di atas selanjutnya dimasukkan dalam rumus:

$$X^2 = \frac{19\left[(6.7 - 2.4) - \frac{19}{2}\right]^2}{(6+8)(4+7)(6+4)(2+7)}$$

$$X^2 = \frac{4,75}{13860} = 0,823$$

   Harga Chi kuadrat tabel untuk dk = 1 dan α 5%(0,05) = 3.841 karena drat hitung lebih kecil dari tabel (0.00034 < 3.841), maka H_0 diterima.

9. Kesimpulan

   Tidak terdapat perbedaan secara signifikan antara penghasilan petani dan nelayan. Berdasarkan mediannya.

15.4.4 Mann-Whitney U-Test

Dalam Uji hipotesis U-test digunakan untuk menguji suatu signifikansi hipotesis komparatif dua sampel independent jika datanya berbentuk ordinal. Pada test ini merupakan jenis test yang memiliki kualitas yang baik untuk menguji suatu hipotesis komparatif dua sampel independent jika, datanya berbentuk ordinal. Jika saat melakukan suatu pengamatan, data berbentuk interval maka perlu diubah terlebih dahulu kedalam bentuk ordinal. Namun jika data masih berbentuk interval, maka dapat menggunakan t-test untuk pengujiannya.

Akan tetapi, test ini tidak dapat digunakan jika asumsi t-test tidak dipenuhi atau misal datanya harus normal. Dalam uji hipotesis U-Test ini, terdapat dua rumus yang biasa digunakan untuk melakukan suatu pengujian. Rumus dibawah ini digunakan untuk mengetahui dan mengidentifikasi mana harga U yang lebih kecil. Tujuan untuk mengetahui harga U yang lebih kecil ini digunakan untuk melakukan suatu pengujian dan membandingkannya dengan U tabel. Berikut rumusnya:

$$U_{1=}n_1n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

Keterangan:

n_1 = jumlah sampel 1

n_2 = jumlah sampel 2

U_1 = jumlah peringkat 1

U_2 = jumlah peringkat 2

R_1 = Jumlah rangking pada sampel n_1

R_2 = Jumlah rangking pada sampel n_2

Kedua rumus diatas digunakan untuk mengetahui dan mengidentifikasi harga U mana yang memiliki harga U paling kecil. Diantara U_1 dan U_2 yang memiliki U paling kecil akan digunakan sebagai U hitung untuk dibandingkan dengan U tabel. Jika nilai U hitung pada U_1 dan U_2 lebih besar dibandingkan dari n_1 n_2/2, maka nilai tersebut adalah nilai , dan nilai U dapat dihitung dengan rumus:

$$U = n_1 \, n_2 - U'$$

Dengan menggunakan kiteria pengambilan keputusan:

a. H0 diterima bila U hitung U tabel (; n_1, n_2)
b. H0 ditolak bila U hitung U tabel (: n_1, n_2)

Contoh:

Suatu penelitian dilakukan untuk mengetahui perbedaan produktivitas petani ladang yang menggunakan teknologi mesin yang canggih dalam memproduksi hasil ladang dengan produktivitas petani ladang yang menggunakan mesin tradisional. Apakah ada perbedaan dalam produktivitas pada petani ladang yang menggunakan mesin tradisional. Berikut diberikan data nilai produktivitas yang diperoleh dari dua kelompok petani ladang tersebut:

TABEL 4.6 ***Produktivitas petani ladang dengan teknologi mesin tradisional dan mesin canggih***

No	Petani ladang (mesin tradisional)	No	Petani ladang (mesin canggih)
	Nilai Produktivitas		Nilai Produktivitas
1.	70	1.	80
2	60	2.	70
3.	50	3.	60
4.	80	4.	80
5.	50	5.	80
6.	70	6.	60
7.	70	7.	70
8.	50	8.	70
9.	80	9.	60
10.	60	10.	80
11.	70	11.	70
12.	60	12.	60
		13	80
		14.	70
		15.	80
		16.	90

Berdasarkan hal tersebut diatas maka:

1. Judul penelitian dapat dirumuskan:
   Perbandingan produktivitas Petani ladang yang menggunakan mesin canggih dan trandisional
2. Peubah penelitiannya adalah:
   a. Produktivitas Petani ladang: peubah independent
   b. Penggunaan jenis mesin
3. Rumusan masalah:
   Apakah ada perbedaan produktivitas petani ladang yang signifikan antara menggunakan mesin canggih dan mesin tradisional.

4. Sampel:
   Terdiri dari 2 kelompok petani yaitu kelompok A (petani ladang dengan mesin tradisional) 12 Petani dan kelompok B (petani ladang dengan mesin canggih) 16 Petani.
5. Kriteria pengujian hipotesis:
   a. H0: Tidak terdapat perbedaan produktivitas petani ladang yang menggunakan mesin tradisional dan mesin teknologi canggih
   b. Ha: terdapat perbedaan produktivitas petani ladang yang menggunakan mesin tradisional dan mesin teknologi canggih
6. Kedua sampel (n_1 dan n_2) diatas digabung untuk dibuat rangking, data gabungan sampel 1(n_1) dan sampel 2 (n_2) disusun dalam tabel seperti ini:

TABEL 4.7 ***Sampel gabungan dengan ranking***

Produktiitas Petani ladang dengan mesin tradisional dan mesin teknologi canggih				
No.	**Nilai Produktivitas (ascending)**	**Sampel gabungan**	**Peringkat/ ranking**	**Jenis gabungan**
1.	70	50	1.	2
2.	60	50	2.	2
3.	50	50	3.	2
4.	80	60	4.	7
5.	50	60	5.	7
6.	70	60	6.	7
7.	70	60	7.	7
8.	50	60	8.	7
9.	80	60	9.	7
10.	60	60	10.	7
11.	70	70	11.	15
12.	60	70	12.	15
13.	80	70	13.	15
14.	70	70	14.	15
15.	60	70	15.	15
16.	80	70	16.	15
17.	80	70	17.	15
18.	60	70	18.	15

Produktiitas Petani ladang dengan mesin tradisional dan mesin teknologi canggih				
No.	Nilai Produktivitas (ascending)	Sampel gabungan	Peringkat/ ranking	Jenis gabungan
19.	70	70	19.	15
20.	70	80	20.	23,5
21.	60	80	21.	23,5
22.	80	80	22.	23,5
23.	70	80	23.	23,5
24.	60	80	24.	23,5
25.	80	80	25.	23,5
26.	70	80	26.	23,5
27.	80	80	27.	23,5
28	90	90	28	28

Tabel diatas selanjutnya dapat ditulis:

TABEL 4.8 ***Produktivitas petani ladang yang menggunakan mesin tradisional dan mesin teknologi canggih dengan rankingnya***

Petani ladang yang menggunakan mesin tradisional			Petani ladang yang menggunakan mesin teknologi canggih		
No.	Nilai Produktivitas	ranking	No.	Nilai Produktivitas	ranking
1.	70	15	1.	80	23,5
2.	60	7	2.	70	15
3.	50	2	3.	60	7
4.	80	23,5	4.	80	23,5
5.	50	2	5.	80	23,5
6.	70	15	6.	60	7
7.	70	15	7.	70	15
8.	50	2	8.	70	15
9.	80	23,5	9.	60	7
10.	60	7	10.	80	23,5
11.	70	15	11.	70	15
12.	60	7	12.	60	7

Petani ladang yang menggunakan mesin tradisional			Petani ladang yang menggunakan mesin teknologi canggih		
No.	**Nilai Produktivitas**	**ranking**	**No.**	**Nilai Produktivitas**	**ranking**
			13	80	23,5
			14.	70	15
			15.	80	23,5
			16.	90	28
JUMLAH		R1 = 134	JUMLAH		R2 = 272

7. Karena Nilai R1 ≤ R2 maka Nilai U dihitung dengan rumus:

$$U_1 \ = n_1 n_2 \ + \frac{n_2(n_2+1)}{2} \ - R_2$$

$$U_1 = 12 \ x \ 16 \ + \frac{16(16+1)}{2} \ - 272$$

$$U_1 = 192 + 136 - 272 = 56$$

Bandingkan dengan hasil berikut:

$$U_2 = n_1 n_2 + \frac{n_1(n_1+1)}{2} \ - R_1$$

$$U_2 = 12 \ x \ 16 + \frac{12(12+1)}{2} \ - 134$$

$$192 + 78 - 134 = 136$$

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1.n_2(n_1+n_2+1)}{12}}}$$

$$Z = \frac{56 - \frac{12.16}{2}}{\sqrt{\frac{12.16(12+16+1)}{12}}} = \frac{-40}{\sqrt{\frac{192(29)}{12}}} = \frac{-40}{\sqrt{\frac{192(29)}{12}}} = \frac{-40}{\sqrt{464}} = \frac{-40}{21{,}540} = -1{,}8570$$

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1.n_2(n_1+n_2+1)}{12}}}$$

$$Z = \frac{136 - \frac{12.16}{2}}{\sqrt{\frac{12.16(12+16+1)}{12}}} = \frac{40}{\sqrt{\frac{192(29)}{12}}} = \frac{40}{\sqrt{\frac{192(29)}{12}}} = \frac{40}{\sqrt{464}} = \frac{40}{21{,}540} = 1{,}8570$$

8. Nilai tabel:
   Pada nilai tabel z, uji dua sisi, α = 5% = 1,96
9. Hasil:
   Menggunakan rumus:
   $U_{hitung} = 1{,}8570 \leq U_{tabel}\ 1{,}96$; Berarti bahwa H0 diterima, dan Ha ditolak
10. Kriteria Pengujian:
   Tidak terdapat perbedaan produktivitas petani ladang yang menggunakan mesin tradisional dan mesin teknologi canggih pada α = 5%.

15.4.5 Test Kolmogorov-Smirnov Dua Sampel

Pengujian hipotesis lainnya pada data berjenis ordinal dapat menggunakan uji Test Kolmogorov-Smirnov. Tetapi sebelum dilakukannya analisis sebaiknya data disusun ke dalam tabel distribusi frekuensi kumulatif mengikut masing-masing kelas interval. Uji statistik pada uji Kolmogorov berfungsi untuk mengetahui ada atau tidaknya terjadi signifikan hipotesis komparatif dua sampel independent. Uji kolmogorov didasarkan oleh distribusi empiris. Keuntungan dari uji Kolmogorov adalah tidak bergantung pada fungsi distribusi kumulatif yang mendasari pengujian. Namun dalam pengujian Kolmogorov juga memiliki kelemahan tersendiri, pada uji ini hanya berlaku untuk distribusi kontinu.

Tujuan dari konsep dasar uji normalitas Kolmogorov adalah dengan membandingkan distribusi data (yang nanti akan diuji normalitasnya) dengan distribusi normal baku. Jika signifikansi dibawah 0,05 maka terdapat perbedaan yang signifikan dan jika terjadi signifikansi diatas 0,05 maka tidak terjadi perbedaan yang signifikan. Pada penerapan uji Kolmogorov, jika signifikansi dibawah 0,05, maka data yang diuji mempunyai perbedaan signifikan dengan data normal baku. Rumus perhitungan uji test ini adalah seperti berikut:

$$D = maksimum\ [Sn_1(X) - Sn_2(X)]$$

Contoh:

Diberi tabel yang berisikan informasi tingkat kesalahan dua mesin produksi makanan di pabrik A. Manajer pabri mencoba untuk meneliti dengan melakukan perbandingan kedua tingkat kesalahan tersebut terhadap produktivitas mesin selama 4 bulan. Setelah melakukan pengamatan data tersaji pada tabel berikut:

Tingkat Kesalahan Mesin Produksi Di Pabrik A (%)

No.	Mesin 1	Mesin 2
1	1,0	6,0
2	2,0	3,0
3	1,0	5,0
4	5,0	7,0
5	5,0	8,0
6	1,0	3,0
7	2,0	4,0
8	1,0	8,0
9	1,0	2,0
10	3,0	5,0

Berdasarkan data di atas, proses penelitian dapat dijalangkan dengan langkah-langkah seperti berikut:

1. Judul yang tepat untuk penelitian di atas adalah perbandingan mesin 1 dan mesin 2 dalam memproduksi makanan di pabrik A.
2. Peubah yang digunakan dalam penelitian adalah kategori mesin sebagai peubah independen dan produktivitas sebagai peubah dependen.
3. Rumusan masalah penelitian tersebut adalah apakah terdapat perbedaan produktivitas kerja yang signifikan antara mesin produksi 1 dan mesin produksi 2
4. Sampel penelitian diambil dari dua jenis mesin yang berbeda namun jumlah kesalahan selama mesin beroperasi adalah sama yaitu sebanyak 10 sampel.
5. Hipotesis untuk penelitian ini adalah:
   Ho: Tidak adanya perbedaan produktivitas yang signifikan antara mesin produksi 1 dan mesin produksi 2
   H1: Ditemukan adanya perbedaan produktivitas yang signifikan antara mesin produksi 1 dan mesin produksi 2
6. Kriteria pemenuhan dalam pengujian hipotesis harus memenuhi syarat yaitu Ho diterima apabila nilai hitung K_D (nilai uji Kolmogorov-Smirnov) lebih kecil atau sama dengan nilai pada tabel (K_D hitung $\leq$ K_D tabel).
7. Perhitungan uji kolmogorov-smirnov
   Data yang tersedia dalam tabel dalam disusun kembali ke dalam tabel distribusi frekuensi kumulatif seperti berikut:

Tabel Interval Tingkat Kesalahan Mesin Produksi 1

No.	Interval	F	Kumulatif
1	1-2	7	7
2	3-4	1	8
3	5-6	2	10
4	7-8	0	10

Kemudian data disusun kedalam tabel seperti di bawah untuk nilai kumulatif dalam bentuk proporsional.

Tabel Penolong untuk Pengujian dengan Kolmogorov-Smirnov

Kelompok	Kesalahan Kerja			
	1-2%	3-4%	5-6%	7-8%
S10 (X)	7/10	1/10	2/10	0/10
S10 (X)	1/10	3/10	3/10	3/10
$Sn_1X - Sn_2X$	6/10	2/10	1/10	3/10

Tabel menunjukkan selisih terbesar pada nilai hitung adalah 6/10. Sehingga, nilai KP_D = 6. Sebelum membandingkan nila KP_D, terlebih dahulu dapatkan nilai tabel K_D dengan pengujian hipotesis satu pihak pada n=10 dan α = 5%. Hasilnya Ho diterima dan H1 ditolak. Kesimpulannya adalah tidak terdapat perbedaan yang signifikan antara produktivitas mesin 1 dan mesin 2. Sedangkan, apabila sampel besan dimana n_1 dan n_2 lebih besar dari 40, pengujian signifikansinya dapat menggunakan rumus perhitungan di bawah yang kemudian membandingkan nilai signifikansi pada tabel K_D dan KP_D. Rumus perhitungan seperti berikut:

$$K_D = 1.36\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

$$K_D = 1.36\sqrt{\frac{10 + 10}{10\,.\,10}} = 0.6$$

Pada contoh diatas harga K_D hitung = 6: 10 = 0,6. Ternyata harga K_D hitung sama dengan harga tabel dengan demikian H_0 tetap diterima (0,6 = 0,6).

8. Kesimpulan dari hasil penelitian di atas adalah tidak adanya perbedaan produktivitas mesin baik mesin produksi 1 atau pun mesin produksi 2.
9. Saran kepada manajer pabrik berdasarkan hasil penelitian adalah penggunaan untuk kedua mesin dapat memberikan hasil yang maksimum baik menggunakan mesin 1 atau pun mesin 2.

15.5 UJI HIPOTESIS KOMPARATIF DUA SAMPEL INDEPENDEN

15.5.1 T-test

Uji T-test digunakan dalam pengujian hipotesis komparatif pada dua sampel denan nilai rata-rata pada data interval atau rasio. Rumus sebagai berikut:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2} - 2r\left(\frac{s_1}{\sqrt{n_1}}\right)\left(\frac{s_2}{\sqrt{n_2}}\right)}}$$

Dimana:

$\overline{x}_1$ = Rata-rata sampel 1

$\overline{x}_2$ = Rata-rata sampel 2

S_1 = Simpangan baku sampel 1

S_2 = Simpangan baku sampel 2

S_1^2 = Varians sampel 1

S_2^2 = Varians sampel 2

r = Korelasi antara dua sampel

Contoh:

Perusahaan ingin melihat apakah dengan penambahan insentif dapat meningkatkan nilai produktivitas para pekerja. Dari 25 sampel pekerja terpilih diperoleh data dikumpulkan secara acak dengan ketentuan pengujian hipotesis seperti berikut.

Ho: Tidak adanya perbedaan nilai produktivitas para pekerja sebelum atau setelah diberi insentif.

H1: Adanya perbedaan nilai produktivitas para pekerja baik sebelum atau setelah diberi insentif.

Nilai Produktivitas 25 Pekerja Sebelum dan Sesudah Diberi Insentif

No. Responden	ProduktivitasKerja	
	Sebelum (x_1)	Sesudah (x_2)
1	65	85
2	75	90
3	70	95
4	80	65
5	65	80
6	80	80
7	65	90
8	70	60
9	90	75
10	70	70
11	60	65
12	70	75
13	75	85
14	70	65
15	80	75
16	75	75
17	80	75
18	65	90
19	70	70
20	75	85
21	75	95
22	80	85
23	80	80
24	90	95
25	75	75
Rata-rata	$\bar{x}_1 = 74$	$\bar{x}_2 = 79.20$
Simpangan Baku	$S_1 = 7.5$	$S_2 = 10.17$
Varians	$S_1^2 = 56.25$	$S_1^2 = 103.5$

Berdasarkan nilai hitung rata-rata produktivitas pekerja sebelum diberi insentif sebesar $\bar{x}_1 = 74$ dengan simpangan baku $s_1 = 7{,}5$, dan varians $S_1^2 = 56.25$. Namun, setelah diberi insenftif nilai $\bar{x}_2$ meningkat kepada 79.20 pada simpangan baku $S_2 = 10.17$, dan varian $S_2^2 = 103.5$. Artinya dengan adanya pemberian insentif maka produktivitas pekerja menjadi meningkat.

Bab 16

Pengujian Hipotesis Asosiatif

16.1 PENGANTAR

Pengujian hipotesis asosiatif berarti menguji hubungan antara dua peubah atau lebih pada sampel penelitian meliputi sebagaian populasi. Fokus utama bab ini adalah untuk menerangkan langkah-langkah pengujian hipotesis asosiatif terhadap tiga jenis data yaitu data nominal, ordinal dan interval. Pengujian akan dilakukan menggunakan tiga pendekatan seperti uji chi kuadrat, uji korelasi spearman, dan uji korelasi pearson. Bab ini juga menjadi bab terakhir dalam buku statistik ini. Oleh karena itu, diharapkan pembaca dapat memahami seluruh konteks yang terdapat dalam setiap bab.

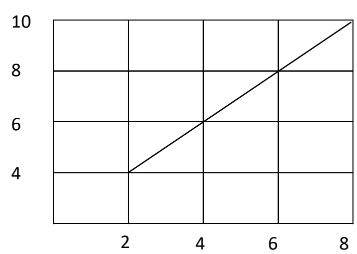
16.2 HIPOTESIS ASOSIATIF

Dalam ilmu statistik, pengujian hipotesis asosiatif perlu untuk dilakukan mengingat ini akan mempengaruhi setiap peubah apakah masing-masing dari peubah mempunyai hubungan satu sama lain. Terdapat tiga bentuk hubungan antar peubah yaitu peubah berhubungan simetris, sebab-akibat (kausal) dan interaktif/resiprocal (saling mempengaruhi). Cara yang dilakukan untuk melihat apakah setiap peubah berpengaruh adalah dengan menghitung korelasi dari setiap peubah. Korealsi ini dalam bentuk angka yang dapat menunjukkan arah kuatnya hubungan masing-masing peubah tersebut. Arah korelasi dinyatakan dalam bentuk positif atau negatif, sedangkan untuk hubungan antar peubah dinyatakan dalam besarnya koefisien korelasi.

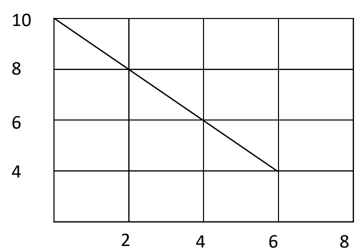
Hubungan antara dua peubah atau lebih apabila didapati positif berarti jika salah satu peubah mengalami peningkatan akan meningkatkan nilai peubah lainnya. Sebaliknya, jika nilai suatu peubah menurun maka akan menurunkan nilai peubah lain. Contoh sederhananya adalah Tinggi badans seseorang terhadap kecepatan berlari. Semakin tinggi badan maka orang tersebut akan semakin cepat berlari. Sebaliknya, semakin pendek badan seseorang maka akan semakin lambat larinya.

Hubungan peubah yang negatif pula menggambarkan hal yang sebaliknya. Apabila nilai suatu peubah mengalami kenaikan maka akan menurunkan nilai pada peubah lain. Sebagai contoh, terdapat hubungan yang negatif antara penjual es dengan curah hujan. Dari contoh tersebut dapat disimpulkan semakin rendah curah hujan maka semakin laris pedagang es. Sebaliknya, semakin tinggi curah hujan maka semakin sedikit jumlah es yang dapat dijual.

Kedua hubungan ini dapat digambar seperti pada Gambar 16.1 berikut:



GAMBAR 16.1 *Korelasi Positif*



GAMBAR 16.2 *Korelasi Negatif*

Pada umumnya, teknik yang digunakan untuk pengujian statistik korelasi adalah uji hipotesis asosiatif. Berikut ini akan lebih dijelaskan jenis analisis apa yang sesuai untuk digunakan mengikut jenis data yang tersedia. Secara ringkasnya, apabila data yang didapat dalam bentuk nominal/diskrit maka teknik yang sesuai untuk pengujian korelasinya adalah menggunakan koefisien kontingensi (Chi Kuadrat), sedangkan untuk data ordinal maka dapat menggunakan uji korelasi spearman rank dan kendal Tau. Terakhir, untuk jenis data interval maka sebaiknya menggunakan uji korelasi pearson dengan koefisien determinasi.

16.3 PENGUJIAN CHI SQUARE

Uji statistik nonparametrik ialah suatu uji statistik yang tidak memerlukan adanya asumsi-asumsi mengenai sebaran data populasinya (belum diketahui sebaran datanya dan tidak perlu berdistribusi normal). Oleh karenanya statistik ini juga dikemukakan sebagai statistik bebas sebaran (tdk mensyaratkan bentuk sebaran parameter populasi, baik normal atau tidak). Statistika non-parametrik dapat digunakan untuk menganalisis data yang berskala Nominal atau Ordinal. Data berjenis Nominal dan Ordinal tidak menyebar normal. Selain itu statistik ini dapat digunakan pada data yang berjumlah kecil, yakni kurang dari 30 data. Dalam pengujian Chi Kuadarat kita dapat menghitung nilai koefisien kontingensi menggunakan rumus berikut:

$$C = \sqrt{\frac{X^2}{N + X^2}}$$

Teknik ini mempunyai hubungan dekat dengan pengujian hipotesis pada hipotesis komparatif k sampel independen. Oleh sebab itu, sangat banyak alternatif untuk pengujian statistik nonparametrik. Berbagai literatur memberikan pengelompokan kategori statistik nonparametrik dengan berbagai cara yang berbeda. Namun demikian, secara sederhana dan berdasarkan prosedur yang sering digunakan, uji-uji tersebut dapat dikelompokkan atas kategori berikut:

- Prosedur untuk data dari sampel tunggal
- Prosedur untuk data dari dua kelompok atau lebih sampel bebas (independent)

- Prosedur untuk data dari dua kelompok atau lebih sampel berhubungan (dependent)
- Korelasi peringkat dan ukuran-ukuran asosiasi lainnya

Distribusi Chi kuadrat digunakan untuk menguji homogenitas varians beberapa populasi. Masih ada beberapa persoalan lain yang dapat diselesaikan dengan mengambil manfaat distribusi chi-kuadrat ini, diantaranya:

1. Menguji proporsi untuk data multinom
2. Menguji kesamaan rata-rata data poisson
3. Menguji independen antara dua faktor didalam kontingensi
4. Menguji kesesuaian antara data hasil pengamatan dengan model distribusi dari mana data itu diduga diambil, dan
5. Menguji model distribusi berdasarkan data hasil pengamatan.

16.3.1 Prosedur Sampel Tunggal dengan Chi-Kuadrat

Akan diuji distribusi frekuensi kategori peubah motivasi hasil amatan dengan distribusi frekuensi kategori peubah sama yang diharapkan. Hipotesis nol uji tersebut adalah: tidak terdapat perbedaan distribusi peubah motivasi hasil amatan dengan distribusi harapan. Prosedur ini banyak digunakan pada uji normalitas peubah.

Misalkan sebuah eksperimen menghasilkan peristiwa-peristiwa atau kategori-kategori $A_1, A_2, \dots, A_k$ yang saling terpisah masing-masing dengan peluang $p_1 = P(A_1), p_2 = P(A_2), \dots, p_k = P(A_k)$.
Akan diuji pasangan hipotesis:

$H_o: p_i = p_{io}, i = 1,2, \dots, k$ dengan p_{io} dengan sebuah harga yang diketahui

$$H_a: p_i \neq p_{io}$$

Disini, tentu saja $\sum p_i = \sum p_{io} = 1$

Agar mudah diingat, adanya kategori A_i, hasil pengamatan O_i dan hasil yang diharapkan E_i, sebaiknya disusun dalam daftar sebagai berikut:

TABEL 16.1 ***Pengelompokan data berdasarkan kelompok***

Kategori	A_1	A_2	…................	A_k
Pengamatan	O_1	O_2	…................	O_k
Diharapkan	E_1	E_2	…................	E_k

Rumus yang digunakan dalam uji tersebut adalah:

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

dengan keterangan:

O_i = Banyaknya kasus yang diamati dalam kategori *i*.

E_i = Banyaknya kasus yang diharapkan

$\sum_{i=1}^{k}$ = Penjumlahan semua kategori *k*.

Contoh:

Dalam suatu eksperimen genetika menurut Mendell telah diketemukan bahwa semacam karakteristik diturunkan meurut perbandingan 1 : 3 : 3 : 9 untuk kategorikategori A, B, C, dan D. Akhir-akhir ini dilakukan 160 kali pengamatan dan terdapat 5 kategori A, 23 kategori B, 32 kategori C dan 100 kategori D. Dengan menggunakan $\propto = 0{,}05$, apakah data di atas menguatkan teori genetika tersebut?

Penyelesaian:

Berdasarkan teori, diharapkan terdapat 1/16 X 160 = 10 kategori A, masing-masing 30 kategori B dan C, dan 90 kategori D. Data hasil pengamatan dan yang diharapkan adalah sebagai berikut.

Kategori	**A**	**B**	**C**	**D**
Pengamatan (O_i)	5	23	32	100
Diharapkan (E_i)	10	20	30	90

Dari rumus didapat:

$$X^2 = \sum_{i=4}^{k} \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(5-10)^2}{10} + \frac{(23-30)^2}{20} + \frac{(32-30)^2}{30} + \frac{(100-90)^2}{90} = 5{,}38$$

Dari tabel distribusi Chi kuadrat diperoleh $X^2_{1-0,05;(4-1)}$ = 7,81. Sehingga pengujian memperlihatkan H_0 diterima yang artinya teori menurut Mendell benar.

Atau dengan cara:

Kategori	O	E	$\frac{(O-E)^2}{E}$
A	5	10	2,5
B	23	30	1,633333
C	32	30	0,133333
D	100	90	1,111111
JUMLAH	160	160	5,377778

Dengan cara tersebut, maka diperoleh X^2 = 5,377 atau 5,38. Derajad kebebasan (db) uji tersebut adalah jumlah kategori (k) dikurangi 1 = 4 – 1 = 3. Pada taraf signifikasi (α) = 5% harga tabel X^2 = 7,81. Karena X^2 hitung < X^2 tabel, maka hipotesis nol diterima.

Latihan:

Diduga bahwa 50% dari semacam kacang bentuknya keriput dan 50% lagi halus. Pengamatan dilakukan terhadap sebuah sampel acak terdiri atas 80 butir kacang dan terdapat 56 keriput sedangkan sisanya halus. Dalam taraf 0,05, dapatkah kita menyokong dugaan tersebut?

Jawaban

Kategori	Pengamatan	Harapan	$\frac{(O - E)^2}{E}$
Keriput	56	40	6,4
Halus	24	40	6,4
	X^{hit}_2		12,8

$$X^2_{tab} = X^2_{0,05;1} = 3{,}841$$

Karena $X^2_{hit} = 12{,}8 > 3{,}84 = X^2_{tab}$ maka H_0 maka ditolak artinya dugaan bahwa 50% kacang keriput dan 50% kacang halus tidak benar

16.3.2 Prosedur untuk Sampel Independen

Hollingshead (1949) meneliti pilihan kurikulum oleh pelajar di kota Elmtown ditinjau dari kelas sosialnya. Kurikulum yang ada mencakup persiapan ke PT, umum, dan perdagangan. Sedangkan kelas sosial yang ada dikelompokkan menjadi 4. Hipotesis nol yang diajukan Hollingshead adalah: proporsi siswa yang tercatat dalam ketiga kemungkinan kurikulum adalah sama untuk semua kelas sosial. Dengan X^2 untuk k sampel independen, rumus yang digunakan adalah sebagai berikut:

$$X^2 = \sum_{i=1}^{r}\sum_{j=1}^{k}\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

dengan keterangan:

O_{ij} = jumlah kasus yang diobservasi dalam baris ke pada kolom ke *j*

E_{ij} = jumlah kasus yang diharapkan dalam baris ke pada kolom ke *j*

$\sum_{i=1}^{r}\sum_{j=1}^{k}$ = jumlah semua sel

Dalam penelitian Hollingshead diperoleh data hasil observasi (O) dan data yang diharapkan (O) seperti Tabel 2. Untuk menghitung $\frac{O_i - E_i}{E_i}$ perlu dibuat kolom $\frac{O_i - E_i}{E_i}$.

TABEL 16.2 *Uji Statistik Nonparametrik Data dari k Sampel Independen dengan Chi-Kuadrat*

Kurikulum	Kelas Sosial						Jumlah O
	I			II			
	E	O	$\frac{(O-E)^2}{E}$	E	O	$\frac{(O-E)^2}{E}$	
Persiapan PT	30.32	40	3.09	38.01	16	12.74	
Umum	77.49	75	0.08	97.13	107	1.00	
Perdagangan	38.18	31	1.35	47.86	60	3.08	
Jumlah		146	4.52		183	16.82	

X^2= 4.52+16.82 =21,34. Derajad kebebasan dalam uji tersebut, db = (2 - 1) * (3 - 1) = 3. Dengan α = 5% dan db = 3 diperoleh $X^2_{tabel} = 0{,}351$.. Karena $X^2_{hitung} > X^2_{tabel}$, maka Ho tidak diterima.

16.3.3 Prosedur untuk Sampel Dependen

Uji McNemar dua sampel dependen dapat diperluas untuk dipakai dalam penelitian yang mempunyai lebih dari dua kelompok sampel. Perluasan ini, yakni uji Q Cochran k sampel berhubungan memberi suatu metode untuk menguji apakah tiga himpunan atau lebih mempunyai frekuensi atau proporsi saling berbeda atau tidak.

Rumus yang digunakan dalam uji Q Cochran adalah:

$$Q = \frac{k(k-1)\left[k\sum_{j=1}^{k} G_j{}^2 - \left(\sum_{j=1}^{k} G_j\right)^2\right]}{k\sum_{i=1}^{N} L_i - \sum_{i=1}^{N} L_i{}^2}$$

dengan keterangan:

G_j = jumlah keseluruhan "sukses" dalam kolom ke *j*

L_I = jumlah keseluruhan "sukses" dalam barir ke *i*.

Misalkan diteliti pengaruh 3 cara wawancara terhadap kemungkinan jawaban dari 10 responden. Jika jawaban pertanyaan "ya" dikode 1 dan jawaban "tidak" dikode 0. Hipotesis nol penelitian ini berbunyi: kemungkinan jawaban "ya" adalah sama untuk ketiga jenis wawancara. Data penelitian ini ditunjukkan pada Tabel 3.

TABEL 16.3. ***Uji Statistik Nonparametrik Data dari k Sampel Dependen dengan Uji Q Cochran***

Responden	Wawancara			L_i^1	L_i^2
	1	2	3		
1	0	0	1	1	1
2	1	0	1	2	4
3	1	1	1	3	9
4	1	1	0	2	4
5	0	1	0	1	1
6	1	0	1	2	4
7	1	1	1	3	9
8	1	0	1	2	4
9	0	0	0	0	0
10	0	0	1	1	1
	$G_1 = 6$	$G_2 = 4$	$G_3 = 7$	$\sum_{i=1}^{10} L_i$	$\sum_{i=1}^{10} L_i{}^2$

Berdasarkan tabel tersebut, maka Q dapat dihitung sbb:

$$Q = \frac{(3-1)\{[3(6)^2 + 3(4)^2 + 3(7)^2] - (17)^2\}}{(3)(17) - 37}$$

$Q = 0.180$

db = k – 1 = 2

$\alpha = 0.05$

Q tabel (X^2tabel) = 5.99.

Q hitung <Q tabel -> Ho diterima.

16.4 PENGUJIAN KORELASI RANK SPEARMAN

Uji korelasi rank spearman dapat digunakan untuk mengetahui hubungan atau pengujian signifikansi hipotesis asosiatif bila masing-masing peubah mempunyai jenis data dalam bentuk ordinal. Selain itu, sumber data yang diperoleh berasal dari sumber yang sama untuk setiap peubah. Metode pengujian ini dapat juga melihat hubungan berdasarkan peringkat (ranking). Sehingga, dengan data berbentuk ordinal atau runut berdasarkan kedudukan maka hasil yang diberikan akan sangat baik dengan ukuran sampel antara 10 hingga 29 sampel. Oleh sebab itu, teknik pengujian ini hanya mengguakan sampel yang kecil. Apabila sampel penelitian melebihi dari jumlah tersebut maka pengujian korelasi rank spearman tidak direkomendasikan. Rumus untuk perhitungan pengujian metode ini adalah sebagai berikut:

$$\rho = 1 - \frac{6 \sum b_i^2}{n\,(n^2 - 1)}$$

ρ = koefisien korelasi
b = selisih pasangan rank ke-i
n = jumlah sampel

Setelah melakukan perhitungan untuk nilai koefisien korelasi dengan rumus di atas, kita dapat menghitung nilai signifikansinya dengan rumus seperti dalam contoh berikut:

Contoh:
Diperoleh hasil pendapat dari 8 orang mahasiswa untuk penilaian terhadap kepuasan di restoran ABC.

Peubah (x)	**Peubah (y)**
3	5
4	3
4	6
3	3
3	2
5	4
4	3
5	5

Jawaban

Berdasarkan dari pengumpulan data seperti yang ditunjukkan dalam tabel di atas, langkah-langkah yang dapat dilakukan untuk pengujian nilai signifikansi korelasi rank spearman sebagai berikut:

1. Menyusun kembali tabel untuk menghitung nilai selisih pada pasangan rank

Peubah (*x*)	**Peubah (*y*)**	***b***	b^2
3	5	-2	4
4	3	1	1
2	6	-4	16
3	3	0	0
3	2	1	1
5	4	1	1
4	3	1	1
5	2	3	9
			$\sum b_i^2 = 33$

2. Menetukan nilai koefisien spearman

$$\rho = 1 - \frac{6 \sum b_i^2}{n\,(n^2 - 1)}$$

$$\rho = 1 - \frac{6\,.\,33}{8\,(64 - 1)}$$

$$\rho = 1 - \frac{198}{504}$$

$$\rho = 0{,}61$$

3. Pengujian nilai signifikansi dapat dilakukan dengan menggunakan rumus berikut:

$$t = \rho \sqrt{\frac{n-2}{1 - \rho^2}}$$

Sehingga, dari contoh sebelumnya kita dapat memasukkan nilai ρ ke dalam rumus di atas.

$$t = 0{,}61 \sqrt{\frac{8-2}{1 - 0{,}61^2}}$$

$$t = 0,61 \sqrt{\frac{6}{0,63}}$$

$$t = 0,61 \,.\, 0,63$$

$$t = 1,88$$

4. Melakukan perbandingan nilai hitung signifikansi di atas (1,88) dengan nilai pada tabel dengan syarat db= 8 – 2 = 6 pada signifikansi 5%. Dengan nilai tabel diperoleh adalah 2,45 maka nilai hitung sebesar 1,88 yang mana lebih kecil dari nilai tabel kesimpulan yang dapat diambil adalah hipotesis Ho diterima sedangkan H1 ditolak.

Contoh:
Dua orang pengamat lukisan menghadiri peluncuran karya-karya lukisan mahasiswa kampus Maju Jaya. Jumlah lukisan yang dikumpulkan adalah 10 buah dan kedua pengamat diminta memberikan penilaian berdasarkan peringkat 1 hinga 10. Nilai yang diberikan dapat dilihat dalam tabel di bawah. Berdasarkan data tersebut tentukan nilai hubungan dari setiap pengamat dalam menilai lukisan yang dipajang. Kesesuain yang diperoleh mengindikasikan penilaian adalah sama.

Jawaban:

1. Judul penelitian yang dapat diberikan untuk penelitian di atas adalah Tingkat kesesuaian pemangat dalam menilai lukisan mahasiswa di kampus Maju Jaya
2. Peubah penelitian terdiri dari dua orang yaitu Pengamat 1 (X1) dan Pengamat 2 (X2)
3. Rumusan masalah yang dapat dibentuk adalah apakah terdapat kesamaan dari kedua pengamat dalam memberikan penilaian terhadap lukisan mahasiswa atau terdapat perbedaan dari hasil penilaian dari pengamat 1 dan pengamat 2?
4. Hipotesis yang dapat diberikan adalah:
   Ho: Tidak adanya kesamaan dari penilaian kedua pengamat dalam memberikan nilai

H1: Didapati adanya kesamaan dari penilaian kedua pengamat dalam memberikan nilai

5. Pengujian hipotesis dapat diterima atau ditolak dengan syarat Ho diterima apabila ρ < nilai tabel ρ
6. Penyajian data dari hasil pengamatan dapat dilihat dalam tabel berikut:

Nilai	Pengamat 1	Pengamat 2
1	6	6
2	7	8
3	8	9
4	2	2
5	3	4
6	4	5
7	7	8
8	5	6
9	6	7
10	9	8

7. Perhitungan untuk pengujian hipotesis pada kasus di atas dapat dilakukan dengan terlebih dahulu merubah data ke dalam bentuk rasio. Perhitungan nilai rasio dapat dilakukan dengan memperhatikan letak angka perolehan setelah diurutkan dari angka terbesar ke terkecil. Dari penilaian para pengamat diperoleh nilai paling besar adalah 9, maka angka ini akan terletak para peringkat 1, angka 8 menjadi peringkat 2. Namun, dalam tabel diperoleh angka 7 pada dua lukisan maka nilai 7 ini berada pada peringkat 3 dan 4 sehingga nilai rasionya harus dibagi dua menjadi (3+4)/2 = 3,5. Jadi, angka 7 dalam tabel perhitungan rasio akan menjadi 3,5 untuk kedua letak. Sebaliknya, pada nilai pengamat 2, angka 8 mempunyai kemunculan yang paling banyak dengan urutan 1, 2 dan 3. Sehingga, nilai letak angka dijumlahkan dan dibagi dua (2+3+4)/3 = 3. Jadi, nilai 8 pada angka tabel akan ditulis 3 mengikut letak angka masing-masing. Terakhir, pada angka 7 akan menjadi angka 5.

Nilai	Pengamat 1	Nilai Rank (X1)	Pengamat 2	Nilai Rank (X2)	*b*	b^2
1	6	5,5	6	6,5	-1	1
2	7	3,5	8	3	0,5	0,25
3	8	2	9	1	1	1
4	2	10	2	10	0	0
5	3	9	4	9	0	0
6	4	8	5	8	0	0
7	7	3,5	8	3	0,5	0,25
8	5	7	6	6,5	0,5	0,25
9	6	5,5	7	5	0,5	0,25
10	9	1	8	3	-2	4
Jumlah					0	7

8. Setelah diperoleh angka b^2 maka kita dapat memasukkan angka tersebut dalam rumus yang telah ditulis seperti berikut:

$$\rho = 1 - \frac{6 \sum b_i^2}{n\,(n^2 - 1)}$$

$$\rho = 1 - \frac{6\ .\ 7}{10\,(10 - 1)}$$

$$\rho = 1 - 0{,}04$$

$$\rho = 0{,}96$$

Berdasarkan angka ρ di atas, kita dapat melakukan perbandingan dengan nilai pada tabel ρ. Untuk sampel sebanyak n = 10, dengan taraf kesalahan 5% diperoleh angka 0,648 sedangkan untuk 1% sebesar 0,794. Berdasarkan hasil ρ tersebut maka Ho ditolak dan H1 diterima karena nilai $\rho >$ nilai tabel ρ.

9. Kesimpulan yang dapat diambil dari hasil pengujian hipotesis di atas adalah terdapat kesesuaian yang signifikan antara pengamat 1 dan pengamat 2 dalam memberikan penilaian kepada 10 lukisan. Keputusan ini diambil berdasarkan pada poin ke-lima yaitu syarat Ho ditolak apabila nilai $\rho <$ dari nilai tabel ρ.

10. Pengujian nilai signifikansi menggunakan rumus *t*.

$$t = \rho \sqrt{\frac{n-2}{1-\ \rho^2}}$$

$$t = 0{,}96 \sqrt{\frac{10-2}{1-\ 0{,}96^2}}$$

$$t = 0{,}96\,.\ 3{,}64$$

$$t = 3{,}49$$

Melakukan perbandingan nilai hitung signifikansi di atas (3,49) dengan nilai pada tabel *t* dengan syarat db= 10 – 2 = 8 pada signifikansi 5%. Dengan nilai tabel diperoleh adalah 2,30 maka nilai *t* hitung sebesar 3,49 yang mana lebih besar dari nilai tabel. Kesimpulan yang dapat diambil adalah hipotesis Ho ditolak dan menerima H1.

16.5 PENGUJIAN KORELASI PEARSON

16.5.1 Metode Korelasi Pearson

Teknik uji validitas item dengan korelasi Pearson yaitu dengan cara mengkorelasikan skor item dengan skor jumlahnya. Skor jumlah adalah penjumlahan seuruh item pada satu peubah. Kemudian pengujian signifikansi dilakukan dengan kriteria menggunakan r tabel pada tingkat signifikansi 0,05 dengan uji 2 sisi. Jika nilai positif dan r hitung ≥ r tabel maka item dapat dinyatakan valid, jika r hitung < r tabel maka item dinyatakan tidak valid

Korelasi merupakan teknik analisis yang termasuk dalam salah satu teknik pengukuran asosiasi atau hubungan (*measure of association*). Pengukuran hubungan merupakan istilah umum yang mengacu pada sekelompok teknik dalam statistika bivariat yang digunakan untuk mengukur kekuatan hubungan antara dua peubah. Diantara sekian banyak teknik-teknik pengukuran korelasi, terdapat dua teknik korelasi yang sangat popular yaitu Korelasi Pearson Product Moment dan Korelasi Rank Spearman. Menurut Hasan (1999:20) menyatakan bahwa, Korelasi Pearson atau sering disebut Korelasi Product Moment (KPM) merupakan alat uji statistik yang digunakan untuk menguji hipotesis asosiatif (uji hubungan) dua peubah bila datanya berskala interval atau rasio. Korelasi Rank Spearman digunakan untuk mencari hubungan atau untuk menguji signifikansi hipotesis asosiatif bila masing-masing peubah yang dihubungkan berbentuk Ordinal (Widarjono, 2013:105).

Analisis korelasi parsial ini digunakan untuk mengetahui kekuatan hubungan anatara korelasi kedua peubah dimana peubah lainnya yang dianggap berpengaruh dikendalikan atau dibuat tetap (sebagai peubah kontrol). Karena peubah yang diteliti adalah data interval maka teknik statistik yang digunakan adalah Pearson Correlation Product Moment (Sugiyono, 2013:216). Rumus untuk Pearson Product Moment adalah sebagai berikut:

$$r_{xy} = \frac{\mathrm{N}\sum \mathrm{XY} - (\mathrm{X})(\mathrm{Y})}{\sqrt{}\mathrm{N}\sum \mathrm{x}\ 2 - \sum \mathrm{X2N}\sum \mathrm{Y2} - \sum \mathrm{Y2}}$$

Keterangan:
rXY = Koefisien Korelasi
N = Banyaknya Sampel
ΣX = Jumlah skor keseluruhan untuk item pertanyaan peubah
XΣY= Jumlah skor keseluruhan untuk item pertanyaan peubah Y

Dari hasil yang diperoleh dengan rumus di atas, dapat diketahui tingkat pengaruh peubah X dan peubah Y. Pada hakikatnya nilai r dapat bervariasi dari -1 hingga +1, atau secara matematis dapat ditulis menjadi $-1 \leq r \leq +1$. Hasil dari perhitungan akan memberikan tiga alternatif, yaitu:

1. Bila r = 0 atau mendekati 0, maka korelasi antar kedua peubah sangat lemah atau tidak terdapat hubungan antara peubah X terhadap peubah Y.
2. Bila r = +1 atau mendekati +1, maka korelasi antar kedua peubah adalah kuat dan searah, dikatakan positif.
3. Bila r = -1 atau mendekati -1, maka korelasi antar kedua peubah adalah kuat dan berlawanan arah, dikatakan negatif.

Menurut Sugiyono (2013:248) penentuan koefisien korelasi dengan menggunakan metode analisis korelasi Pearson Product Moment dengan menggunakan rumus sebagai berikut:

$$r_{xy} = \frac{\text{n} \sum \text{xiyi} - (\sum \text{xi})(\sum \text{yi})}{\sqrt{\{\text{n} \sum \text{xi 2} - (\sum \text{xi) 2}\} - \{\text{n} \sum \text{yi 2} - (\sum \text{yi) 2}\}}}$$

Terdapat dua dari beberapa teknik korelasi yang sangat populer sampai sekarang yaitu Korelasi Pearson Product Moment dan Korelasi Rank Spearman. Korelasi Pearson merupakan korelasi sederhana yang hanyamelibatkan satu peubah terikat (dependent) dan satu peubah bebas (independent). Korelasi Pearson menghasilkan koefesien korelasi yang berfungsi untuk mengukur kekuatan hubungan linier antara dua peubah. Jika hubungan dua peubah tidak linier, maka koefesien korelasi Pearson tersebut tidak mencerminkan kekuatan hubungan dua peubah yang sedang diteliti, meski kedua peubah mempunyai hubungan kuat. Koefisien korelasi ini disebut koefisien korelasi Pearson karena diperkenalkan pertama kali oleh Karl Pearson tahun 1990 (Firdaus, 2009). Seperti yang diungkapkan oleh Ronny Kountur (2009:210) bahwa data yang berskala interval atau rasio dapat menggunakan korelasi Pearson.

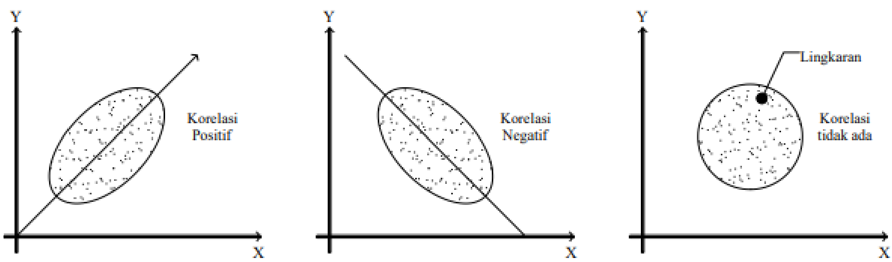
Selain itu, signifikansinya tidak hanya harus memenuhi persyaratan pengukuran tersebut, tetapi harus pula menganggap data berdistribusi normal. Simbol untuk korelasi Pearson adalah “p” jika diukur dalam populasi, dan “r” jika di ukur dalam sampel. Koefisien korelasi adalah ukuran yang dipakai untuk mengetahui derajat hubungan antara peubah-peubah (Siregar, 2013). Nilai koefesien korelasi berada di antara -1<0.

Salah satu teknik satatistik yang kerap kali digunakan untuk mencari hubungan antara dua peubah atau lebih adalah teknik korelasi. Dua peubah yang hendak diselidiki hubungannya tersebut biasanya diberi simbol peubah X dan peubah Y. Bila mana kenaikan nilai peubah X selalu disertai kenaikan peubah Y, dan turunnya nilai peubah X juga selalu diikuti oleh turunnya nilai peubah Y, maka hubungan yang seperti itu disebut hubungan yang positif. Akan tetapi, sebaliknya bilamana kenaikan nilai peubah X selalu diikuti oleh penurunan nilai peubah Y, dan penurunan nilai peubah X justru diikuti oleh kenaikan nilai peubah Y, maka hubungan antara peubah X dan Y tersebut adalah hubungan yang negatif. Disamping itu, dua peubah X dan Y ada kemungkinannya tidak memiliki hubungan sama sekali, yakni bilamana kenaikan nilai peubah yang satu kadang-kadang diikuti penurunan nilai peubah lainnya, dan kadang-kadang juga diikuti oleh kenaik Disamping itu, untuk menafsirkan harga r (koefisien korelasi) maka dapat dikonsultasikan (dibandingkan) dengan harga kritik r product moment (tabel r). an nilai peubah yang lainya.

16.5.2 Koefisien Hubungan

Pada umumnya besar kecilnya hubungan dinyatakan dengan bilangan. Bilangan yang menytatakan besar kecilnya hubungan tersebut disebut koefisien hubungan atu koefisien korelasi. Koefisien korelasi itu berkisar antara 0,00 dan +1,00 (korelasi positif) dan atau diantara 0,00 sampai -1,00 (korelasi negatif), tergantung pada arah hubungan positif ataukah negatif. Koefisien yang bertanda positif menunjukkan bahwa arah korelasi tersebut positif, dan koefisien yang bertanda negatif menunjukkan arah korelasi yang negatif. Sedangkan koefisien yang bernilai 0,00 menunjukkan tidak adanya korelasi antara peubah X dan Y. Bila mana dua peubah mempunyai koefisien korelasi sebesar +1,00 maka berarti bahwa dua peubah tersebut mempunyai korelasi positif yang sempurna. Sebaliknya bilamana dua peubah mempunyai koefisien korelasi -1,00, maka berarti dua peubah tersebut memiliki korelasi negatif yang sempurna. Korelasi yang sempurna semacam itu sangat jarang sekali dijumpai dalam praktik penyelidikan/penelitian. Korelasi antara dua peubah pada umumnya akan berkisar antara +1,00 sampai dengan -1,00.

ILUSTRASI:



16.5.3 Korelasi Product Moment

Untuk menerapkan koefisien korelasi antara dua peubah yang masing-masing mempunyai skala pengukuran interval maka digunakan korelasi product moment yang dikembangkan oleh Karl Pearson. Rumus korelasi product momen ini ada dua macam, yaitu:

1. Korelasi product moment dengan rumus simpangan (deviasi).
2. Korelasi Product moment dengan rumus angka kasar. Korelasi product moment dengan rumus simpangan (deviasi).

$$r_{\varkappa y} = \frac{\sum \varkappa . y}{\sqrt{(\sum \varkappa\ 2)(\sum y 2)}}$$

Dalam hal ini:

$r\varkappa y$ = Koefisien korelasi antara peubah X dan Y.
$\varkappa$ = deviasi dari mean untuk nilai peubah X
y = deviasi dari mean untuk nilai peubah Y
$\sum \varkappa . y$ = jumlah perkalian antara nilai X dan Y
$\varkappa_2$ = Kuadrat dari nilai $\varkappa$
y_2 = Kuadrat dari nilai y

16.5.4 Intrerpretasi Harga

r Interpretasi terhadap harga atau koefisien korelasi secara konvensional diberikan oleh Guilford (1956) sebagai berikut:

Koefisien korelasi r	Interpretasi
0,80 – 1,00	Sangat Tinggi
0,60 – 0,80	Tinggi
0,40 – 0,60	Cukup
0,20 – 0,40	Rendah
0,00 – 0,20	Sangat Rendah

Disamping itu, untuk menafsirkan harga r (koefisien korelasi) maka dapat dikonsultasikan (dibandingkan) dengan harga kritik r product moment (tabel r). Dalam hal ini, ditentukan tingkat kesalahan (peluang ralat) adalah 5% (yang biasa digunakan pada ilmu-ilmu social) dengan melihat pada tabel r berdasarkan N= banyaknya responden. Contoh: pada perhitungan korelasi product moment dimuka diperoleh harga r=0,745.

Harga r kritik (r tabel) pada tingkat kesalahan 5% dan N=10 adalah r tab=0,632. Berarti harga r yang diperoleh dari perhitungan (r hitung)=0,745 > r tab= 0,632. Hal ini menunjukkan bahwa korelasi antara dua peubah tersebut berarti (signifikan). Jika r hitung ternyata < r tabel maka dikatakan bahwa korelasi antara kedua peubah tersebut tidak berarti (tidak signifikan). Jadi, meskipun ada korelasi tetapi secara statistic kurang berarti.

16.5.5 Uji Signifikansi r

Untuk menuji signifikansi koefisien korelasi (nilai r) yang diperoleh maka dapat dilakukan sebagai berikut:

1. Dengan mengacu pada criteria koefisien korelasi yang diberikan oleh Guilford (1956).
2. Dengan membandingkan nilai r hitung dengan harga r tabel dengan taraf kesalahan (α=0,05) atau α=0,01 dan db=N-2.
3. Dengan menghitung lebih dulu t hitung berdasarkan harga r hitung yang diperoleh, yakni dengan rumus sebagai berikut:

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

Untuk korelasi Product Moment, harga t hitung yang diperoleh selanjutnya dikonsultasikan dengan harga t tabel dengan taraf signifikansi tertentu (misal: α=0,05 atau α=0,01) dan dengan derajad kebebasan dk=N-2.

Bila nilai t hitung > t tabel maka tolak Ho, dan berarti menerima H1. Sedangkan bila nilai t hitung < t tabel, maka tidak menolak Ho, yang berarti menolak H1.

Contoh:
Peneliti dalam bidang kesehatan sedang melakukan pengujian dalam melihat apakah tekanan sistolik mempunyai hubungan dengan usia. Dari jumlah populasi yang besar ditetapkan sampel secara acak sebanyak 7 orang dengan rincian data seperti usia, berat badan dan tekanan sistoliknya. Berdasarkan data yang disajikan dalam tabel hitung nilai koefisien korelasi dan lakukan pengujian hipotesis dengan tingkat signifikasi sebesar 5% serta berikan kesimpulan dari perolehan kedua perhitungan. Berikut adalah data sampel tersebut:

Responden	Berat Badan	Usia	Tekanan Sistolik
1	44	43	129
2	45	34	108
3	57	58	149
4	55	78	174
5	56	49	126
6	65	64	168
7	63	73	161

Jawaban:

1. Dapatkan nilai X dan Y kuadrat serta perkalian antara peubah X dan Y.

Responden	Usia (X)	Tekanan Sistolik (Y)	X^2	Y^2	X . Y
1	43	129	1849	16641	5547
2	34	108	1156	11664	3672
3	58	149	3364	22201	8642
4	78	174	6084	30276	13572
5	49	126	2401	15876	6174
6	64	168	4096	28224	10752
7	73	161	5329	25921	11753
Jumlah	399	1015	24279	150803	60112

2. Hitung nilai koefisien korelasi dengan menggunakan rumus.

$$r_{xy} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \ . \ \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

$$r_{xy} = \frac{7\,(60112) - (399)(1015)}{\sqrt{7\,(24279) - (399)^2} \ . \ \sqrt{7\,(150803) - (1015)^2}}$$

$$r_{xy} = \frac{15799}{103{,}69 \ . \ 159{,}36} = 0{,}9561$$

3. Setelah memperoleh nilai dari koefisien korelasi, maka kita dapat melakukan pengujian hipotesis menggunakan rumus t di atas.

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

$$t = 0{,}9651 \cdot \frac{\sqrt{7-2}}{\sqrt{1-(0{,}9651)^2}}$$

$$t = 7{,}30$$

5. Tentukan nilai t-tabel dan lakukan perbandingan.
   Berdasarkan jumlah sampel, maka dk = 7 – 2 = 5. Kemudian, dengan nilai t-tabel pada $\alpha = 5\%$ atau 0,05 sehingga kita dapati nilai tabel t adalah sebesar $t_{(0{,}05/2,\ 5)} = 2{,}57$.
   Perbandingan nilai t yang diperoleh yaitu sebesar 7,30 dan nilai pada t-tabel terletak pada nilai 2,57. Dengan nilai $t >$ daripada nilai t-tabel maka kesimpulannya adalah Ho ditolak dan H1 diterima.
6. Kesimpulan atas interpretasi nilai perbandingan dan keputusan pengujian hipotesis.
   Dari hasil perbandingan pada langkah 5 di atas, dimana Ho ditolak dan H1 diterima maka dapat disimpulkan bahwa terdapat hubungan linear antara usia dan tekanan sistolik.
7. Tentukan nilai koefisien determinasi.
   Nilai r pada koefisien korelasi adalah sebesar 0,9651. Dengan memasukkan nilai koefisien ini pada rumus koefisien determinasi maka kita akan memperoleh nilai seperti berikut:

$$Koefisien\ Determinasi\ (R) = r^2$$

$$Koefisien\ Determinasi\ (R) = (0{,}9651)^2$$

$$Koefisien\ Determinasi\ (R) = 0{,}91$$

Dari nilai koefisien determinasi di atas, maka dapat dinyatakan bahwa keragaman pada proporsi tekanan sistolik dapat dijelaskan keduanya mempunyai hubungan satu sama lain. Dengan standar error atau nilai signifikan berada pada tingkat 5% atau diyakini 95% dapat diterima bahwa sebanyak 91% keragaman tekanan sistolik seseorang ditentukan oleh usia.

Daftar Pustaka

Amin. I., Aswin. A., Fajar. I., Isnaeni, Iwan. S., Pudjirahaju. A., & Sunindya. R. 2009. Statistika untuk Praktisi Kesehatan. Yogyakarta: Graha Ilmu.

Black, J. A., & Dean, J. C. 1999. Metode dan Masalah Penelitian Sosial. Bandung: Refika Aditama.

Dahlan. S. M. 2012. Statistik untuk Kedokteran dan Kesehatan. Jakarta. Salemba Medika.

Hadi, I. 2019. Statistik II untuk Ilmu Sosial dan Ekonomi. Cetakan Kesatu. Surabaya: UWKS Press.

Hadi, S. 2018. Statistik Inferensial, Teori dan Aplikasinya. Edisi Kedua. Banjarmasin: Rajawali Pers.

Hadi. S. 2002. Statistik. Jilid 2. Yogyakarta: Andi Offset.

Hamzah, L. M., Awaludin, I., Maimunah, E. 2016. Pengantar Statistika Ekonomi. Bandar Lampung: CV. Anugrah Utama Raharja (AURA).

Hasibuan. A. A., Supardi, S. D. 2009. Pengantar Statistik Pendidikan. Jakarta: Gaung Persada Press.

Heryana, A. 2016. Uji Statistik Non Parametrik. Jakarta: Fakultas Ilmu-Ilmu Kesehatan, Universitas Esa Unggul.

Nuryadi, Astuti, T. D., Utami, E. S., & Budiantara, M. 2017. Dasar-Dasar Statistik Penelitian. Yogyakarta: Sibuku Media.

Purnomo, R. A. 2017. Analisis Statistik Ekonomi dan Bisnis dengan SPSS. Cetakan Ketiga. Ponorogo: UNMUH Ponorogo Press.

Radiasari. 2014. Statistik Non Parametrik. Malang: Universitas Brawijaya.

Riduwan. 2010. Dasar-dasar Statistika. Bandung: Alfabeta.

Riwidikdo, H. 2012. Statistik Kesehatan. Yogyakarta: Nuha Medika.

Setiabudi, N. A. 2011. Analisis Data Kategorik (STK351). Bogor: Departemen Statistika, FMIPA IPB.

Siegel, S. 1999. Statistik Nonparametrik untuk Ilmu-ilmu Sosial. Jakarta: Gramedia.

Singarimbun, M. & Effendi, S. 1995. Metode Penelitian Survai. Jakarta: LP3ES.

Siswandari. 2009. Statistika (Komputer Based). Surakarta: LPP UNS dan UNS Press.

Suciptawati, N. L. P. 2016. Penuntun Praktikum Statistika Non Parametrik Dengan SPSS 21. Bukit Jimbaran: Fakultas MIPA,Universitas Udayana.

Sudjana. 1989. Metoda Statistika. Bandung: Penerbit Transito.

Sugiyono. 2001. Statistik Nonparametrik untuk Penelitian. Bandung: Alfabeta.

Sugiyono. 2003. Statistika untuk Penelitian. Bandung: Alfabeta.

Sugiyono. 2009. Statistik Non Parametris Untuk Penelitian. Bandung: Alfabeta.

Suyanto, & Gio, P. U. 2017. Statistika Nonparametrik dengan SPSS, Minitab, dan R. Medan: USU Press.

Syarifuddin, A. 2016. Buku Ajar Statistika Nonparametrik. Nusa Tenggara Barat: Fakultas Pertanian, Universitas Mataram.

Wandasari, N. 2014. Uji 1 Sampel Bag 1a (Uji Binomial). Jakarta: Program Studi Kesehatan Masyarakat, Universitas Esa Unggul.

Wibisono, Y. 2005, Metode statistik. Yogyakarta: Gadjah Mada University Press.

Wirawan, N. 2016. Cara Mudah Memahami Statistika Ekonomi dan Bisnis (Statistik Deskriptif). Edisi Keempat. Denpasar: Keraras Emas.

Sudjana. 2005. Metode Statistika. Bandung: Tarsito.

Johnson, R. A. 2010. Statistics: Principles and Methods (6rd ed). Hoboken, NJ: John Wiley & Sons .

Howell, D. C. 2011. Fundamental Statistics for the Behavioral Sciences (7rd ed). Belmont, CA: Wadsworth.

Arief, F. 2011. Pengantar Penelitian dalam Pendidikan. Yogyakarta: Pustaka Pelajar.

Tentang Penulis

Ir. S. Benny Pasaribu. M.Ec., Ph.D

Sahala Benny Pasaribu lahir di Medan, 21 Desember 1958. Tahun 1976, Benny diterima sebagai mahasiswa di Intitut Pertanian Bogor. Pada tahun 1978 Benny terpilih menjadi Ketua Umum Himpunan MISETA (Perhimpunan Mahasiswa Peminat Ilmu Sosial Ekonomi Pertanian). Lulus dari IPB tahun 1981, Benny kemudian diterima sebagai Staf Khusus Menteri Koperasi/ Kepala Bulog. Setelah memperoleh banyak pengalaman di Kementerian, pada tahun 1988 Benny mendapatkan beasiswa dari *United States Agency for International Development* (USAID), U.S.A. dan melanjutkan pendidikannya ke jenjang Strata 2 bidang Ilmu Ekonomi Pembangunan di Williams College, Massachusetts, Amerika Serikat. Setelah menyelesaikan pendidikan S2, Benny mendapatkan beasiswa dari *Canadian International Development Agency* (CIDA), Canada dan melanjutkan program Doktoralnya atau S-3 bidang Ekonomi Industri (Industrial Organization) dan Perdagangan Internasional (*International Trade*) di Ottawa University, Kanada (1990). Bahkan selama berstatus sebagai mahasiswa, Benny diamanahkan menjadi Ketua Umum Perhimpunan Mahasiswa Indonesia di Kanada.

Sepanjang perjalanan karirnya, Benny pernah menjabat pada berbagai sektor pemerintahan yaitu, (1) Kepala Biro Perencanaan Departemen Koperasi dan Pembinaan Pengusaha Kecil RI (1997-1998), (2) Deputi Menteri/Kepala Badan Penanaman Modal dan Pembinaan BUMN Bidang Usaha Sektor Pertambangan, Energi, Industri, dan Telekomunikasi (1999-2000), (3) Ketua Komisi IX DPR RI Membidangi Keuangan, Perbankan, BUMN, dan Perencanaan (2000-2004), (4) Ketua Panitia Anggaran DPR RI (2000-2004), dan (5) Ketua Komisi Pengawas Persaingan Usaha (KPPU) RI selama dua periode (2006-2013). Benny juga diangkat oleh Presiden RI sebagai Anggota Komite Ekonomi dan Industri Nasional (KEIN) RI Periode tahun 2016–2019, dan dipilih sebagai Ketua Pokja Penyusunan *Roadmap* Industrialisasi Indonesia 2045. Selain dibidang pemerintahan, Benny juga aktif di Organisasi Masyarakat dan Profesi, seperti Pengurus Harian HKTI dan menjadi Waketum KADIN Indonesia.

Saat ini, Benny aktif sebagai Dosen dan sekaligus Ketua Senat di Universitas Trilogi. Sebelumnya Benny juga aktif mengajar di Ottawa University, Canada, Universitas Indonesia, Universitas Gajah Mada, dan STIE Perbanas. Adapun bidang yang ditekuninya yaitu: Mikroekonomi, Makroekonomi, Matematika Ekonomi, Ekonomi Pembangunan, Ekonomi Industri, Ekonomi Regulasi, Managemen Produksi, dan Managemen Pemasaran. Saat ini Benny juga aktif menjadi Narasumber Seminar pada lini Pendidikan Tinggi dan Pemerintahan. Benny juga aktif menulis di berbagai media massa dan memberikan masukan pada beberapa kebijakan Pemerintah.

Rizqon Halal Syah Aji. S.Si., M.Si., Ph.D
Adalah Dosen tetap Fakultas Ekonomi dan Bisnis Universitas Islam Negeri Syarif Hidayatullah Jakarta. Lahir di Brebes, pada 5 April 1979 dari Ayah KH. Tajuddin Noor dan Ibu Hj. Toinah. Bergabung dengan UIN Jakarta pada tahun 2011 dan bertugas mengampu mata kuliah rumpun ilmu kuantitatif pada program studi Ekonomi Islam. Kemudian pada tahun 2015

mutasi ke Fakultas Ekonomi dan Bisnis dan bergabung dengan Program Studi Ekonomi Pembangunan. Pada tahun 2016, Rizqon diberi tanggung jawab sebagi sekretaris program Studi Ekonomi Pembangunan. Jenjang pendidikan formal Strata-1 ditempuhnya di Universitas Islam As-Syafi'iyah Jakarta sebagai Mahasiswa Jurusan Matematika Terapan dengan konsentrasi Statistika terapan dan pada akhir studinya ia dikukuhkan Menjadi Wisudawan Terbaik Pertama tingkat Universitas pada tahun 2005. Ketika duduk sebagai mahasiswa matematika (2003), Ia memegang amanah sebagai Ketua Badan Pekerja Ikatan Himpunan Mahasiswa Statistika Indonesia (IHMSI). Perjalanan pendidikan formalnya dilanjutkan pada program Magister Kajian Kependudukan dan Ketenagakerjaan di Universitas Indonesia dan lulus pada tahun 2009 di bawah bimbingan pakar Ekonometrik Indonesia yakni Prof. Dr. I Gusti Ngurah Agung. Seiring dengan kesibukannya di perkuliahan tingkat Magister, Ia banyak berguru seputar riset kuantitatif khususnya untuk survei dan jejak pendapat umum di Lembaga Penelitian, Pendidikan dan Penerangan Ekonomi Sosial (LP3ES), di bawah mentor langsung pioner Survei Politik Indonesia yakni KH. Enceng Shobirin Nadj sekaligus sebagai Direktur *Center for the Study of Development and Democracy (CESDA).* Pada tahun 2005, ia juga merintis lembaga survei bersama mentornya yang diberi nama Survei *Swamedia Research & Communication (SRC),* hingga sang mentor wafat (2020). Ia meneruskan pendidikan pada tingkat Doktoral (2017) di Universiti Kebangsaan Malaysia (UKM) pada disiplin Ilmu Ekonomi dan Lulus pada tahun 2020. Pada Tahun 2020 juga, Rizqon bergabung dengan Yayasan Pengembangan Pendidikan Indonesia Jakarta (YPPIJ/Universitas Trilogi) sebagai staf khusus Yayasan di bawah kepemimpinan Prof. Dr. Ir. Arissetyanto Nugroho, sekaligus ditugaskan untuk mengajar mata kuliah metodologi penelitian pada program studi Magister Manajemen. Pengalaman mengajar dimulai dari almamaternya yakni Universitas Islam As-syafiiyah (UIA), Kemudian mengajar di Universitas Islam Azahra, serta Sekolah Tinggi Agama Islam Al-Ayubi. Banyak karya-karya ilmiah yang telah ditorehnya dalam bidang Ilmu Ekonomi maupun sosial pada jurnal Internasional maupun jurnal nasional bereputasi. Karya terpopuler yang Ia tulis adalah buku dengan judul; **Ansor dan Tantangan Kebangsaan, Sebuah Refleksi Demografi Politik Dari *Social Capital* Menuju *Human Capital*** yang diterbitkan oleh Penerbit Buku Republika (2015).

Dr. Kabul Wahyu Utomo. SE., M.Si

Kabul lahir di Kebumen, Jawa tengah pada 13 Juli 1975. Kabul adalah anak sulung dari 4 bersaudara. Masa kecilnya sampai SMA dihabiskan di Kebumen. Kabul menempuh pendidikan di SD Inpres/SD Negeri 3 Pejagoan, SMP Negeri 1 dan SMA Negeri 1 Kebumen. Tahun 1993, Kabul melanjutkan studinya selama 3,5 tahun di Jurusan Manajemen, FEB Universitas Islam Indonesia, Yogyakarta. Lulus S1, Kabul belajar bisnis, manajemen dan pengelolaan perusahaan dari Ayahnya. Pada tahun 1999, panggilan hati untuk terus belajar membawanya studi S2 di Program Magister Sains, Program Studi Manajemen, Universitas Gadjah Mada Yogyakarta. Setelah menyelesaikan pendidikan S2 pada tahun 2001, Kabul memulai karirnya sebagai dosen dan dipercaya mengajar di beberapa Perguruan Tinggi Swasta di Yogyakarta dan kota-kota di sekitarnya. Kabul menjadi dosen tetap di STIE Putra Bangsa Kebumen (sekarang Universitas Putra Bangsa Kebumen). Kemudian, pada akhir 2003, Kabul memutuskan untuk memulai karirnya di Jakarta dengan menjadi dosen di Sekolah Tinggi Ilmu Ekonomi Keuangan dan Perbankan Indonesia (STEKPI). Saat mengabdi sebagai Dosen Universitas Trilogi (d/h STEKPI). Pada tahun 2016 Kabul melanjutkan studinya dan berhasil meraih gelar Doktor dengan menyelesaikan Pendidikan Doktor Manajemen di Program Doktor, Fakultas Ekonomika dan Bisnis Universitas Gadjah Mada Yogyakarta.

Kabul juga diamanahkan untuk mengemban beberapa jabatan struktural di kampus sebagai Wakil Ketua Bidang Akademik STIE Putra Bangsa Kebumen (Sekarang Universitas Putra Bangsa Kebumen) tahun 2001. Kemudian menjadi Kasubag pada tahun 2004. Di Universitas Trilogi (d/h STEKPI) menjabat sebagai Pelayanan Pembelajaran. Setelah tiga bulan kemudian, Kabul dipromosikan menjadi Kabag. Humas dan Kerjasama merangkap jabatan sebagai Kabag. kemahasiswaan dan Perpustakaan.

Oktober 2016, Kabul dipercaya menjadi Wakil Direktur Magister Manajemen di Universitas Trilogi. Tiga bulan kemudian Kabul diangkat menjadi Wakil Rektor Bidang Akademik dan Kemahasiswaan Universitas Trilogi menggantikan pejabat sebelumnya hingga akhir masa jabatan selesai.

Selesai masa jabatan Wakil Rektor, amanah baru diberikan kepada Kabul untuk menjadi Ketua Prodi program Magister Manajemen merangkap Ketua Program Studi Manajemen (S1) Universitas Trilogi dan mengawal proses Akreditasi S2-MM hingga selesai.

Pada Tahun 2019, Kabul diangkat menjadi Dekan Fakultas Ekonomi dan Bisnis merangkap sebagai Dekan Fakultas Bio Industri Universitas Trilogi. Setelah menjabat sebagai Dekan selama 3 bulan, Kabul kembali diamanahkan menjadi Wakil Rektor Bidang Akademik dan Kemahasiswaan Universitas Trilogi hingga saat ini. Selain aktif sebagai dosen, Kabul juga aktif meneliti dan mengembangkan konsep VIEWS (*Value, Integrity, Entrepreneurship, Wisdom,* dan *Sincerely*) dan menjadi narasumber di beberapa seminar, workshop, maupun webinar.***

DR. Aty Herawati

Aty Herawati lahir di Bandung 26 Februari 1970. Menyelesaikan Sarjana di jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Padjadjara Bandung pada tahun 1993 dan melanjutkan Magister di Universitas Indonesia pada tahun 2000. Pendidikan terakhir ditempuh di Program Doktor Manajemen Bisnis Institut Pertanian Bogor, lulus tahun 2016 dengan konsentrasi Manajemen Keuangan.

Memulai Karir sebagai Dosen yang Dipekerjakan di Universitas Mercu Buana dengan jabatan terakhir sebagai Ketua Program Studi Magister Manajemen Program Pasca Sarjana, saat ini Aty Herawati berkarir sebagai Dosen yang Diperkerjakan di Universitas Trilogi dengan mengemban amanah sebagai Kepala Lembaga Penelitian dan Pengabdian Kepada Masyarakat.