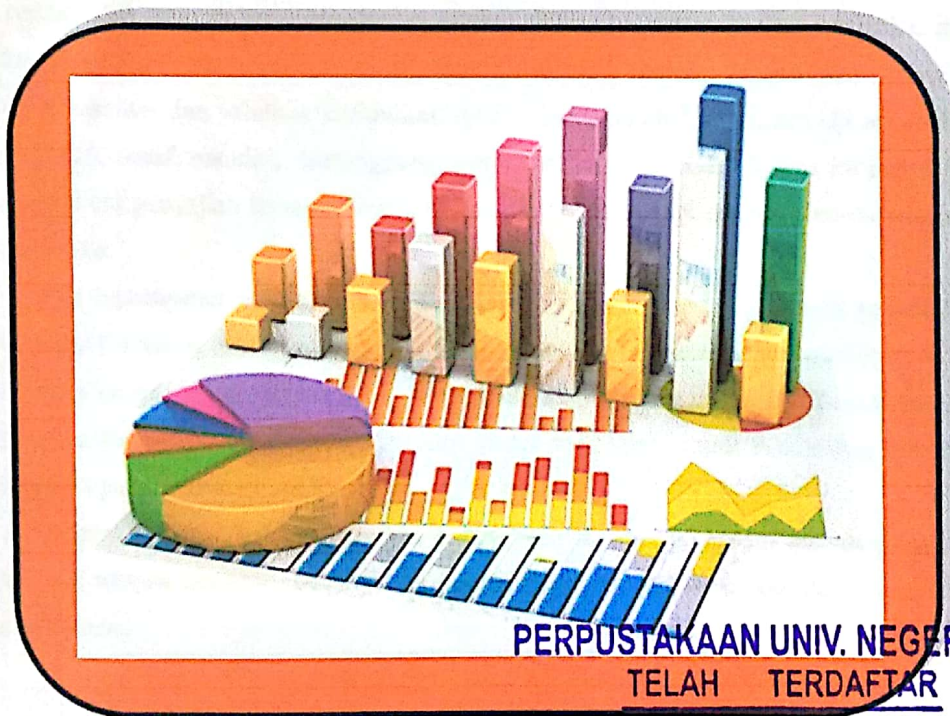


# STATISTIKA

*Modul untuk Mahasiswa Pendidikan Teknik Informatika*

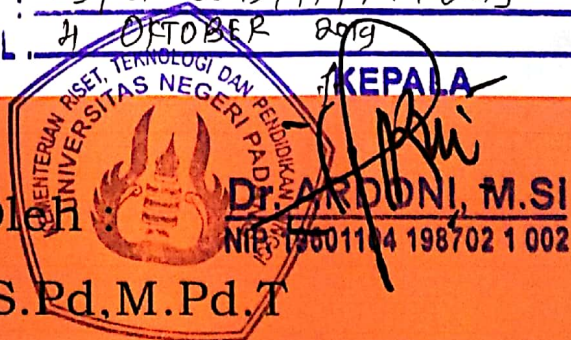


PERPUSTAKAAN UNIV. NEGERI PADANG  
TELAH TERDAFTAR

JUDUL : STATISTIKA  
PENGARANG : IKA PARMA DEWI, S.Pd, M.Pd.T  
JENIS : MODUL  
NOMOR : 69/UN.35.15/PF/F1/2019  
TANGGAL : 4 OKTOBER 2019

Disusun Oleh

Ika Parma Dewi, S.Pd, M.Pd.T



# STATISTIKA

*Modul untuk Mahasiswa Pendidikan Teknik Informatika*



Disusun Oleh :

Ika Parma Dewi, S.Pd, M.Pd.T

## KATA PENGANTAR

Syukur alhamdulillah kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayahNya sehingga penulis dapat menyelesaikan dan menghadirkan modul “Statistika” ini. Modul ini diterbitkan sebagai salah satu tugas akhir pada mata kuliah Landasan Pendidikan dan Pembelajaran di Program Studi Pendidikan informatika. Modul ini merupakan bahan ajar bagi mahasiswa khususnya jurusan Pendidikan informatika agar memahami tentang Berpikir kritis, cerdas, dan teliti dalam menjelaskan konsep dasar dan ruang lingkup statistika, metode statistika, klasifikasi dan distribusi data, ukuran tendensi sentral, ukuran penyebaran data, hipotesis, normalitas dan validitas, reliabilitas, uji korelasi, , uji chi <sup>2</sup>, Uji t, dan Uji anova dengan jujur, kreatif, inovatif, mandiri, bertanggung jawab, dan religius .

Pada kesempatan kali ini penulis juga ingin mengucapkan terima kasih kepada Bapak Drs. Efrizon, MT sebagai dosen pembina mata kuliah yang telah memberikan masukan, motivasi dan dukungan kepada penulis untuk dapat menyelesaikan penulisan buku ini. Terima kasih juga kepada bapak/ibu dosen dan semua rekan-rekan seangkatan di Program Studi Pendidikan Informatika atas kerjasama yang begitu kompak.

Dengan berbagai kekurangan dan kelemahannya, semoga modul ini akan membawa manfaat yang sebesar-besarnya untuk proses pembelajaran dan dalam rangka ikut mencerdaskan bangsa Indonesia.

## SILABUS

### RANCANGAN PEMBELAJARAN SATU SEMESTER

---

**Nama Mata Kuliah** : Statistika  
**Program Studi** : Pendidikan Teknik Informatika-S1  
**Fakultas** : Teknik UNP  
**Dosen** : Ika Parma Dewi, S.Pd., M.Pd.T.  
**SKS** : 2 (dua)

---

Learning Outcomes (Capaian Perkuliahan) Mata kuliah Terkait KKNi

Berpikir kritis, cerdas, dan teliti dalam menjelaskan konsep dasar dan ruang lingkup statistika, metode statistika, klasifikasi dan distribusi data, ukuran tendensi sentral, ukuran penyebaran data, hipotesis, normalitas dan validitas, reliabilitas, uji korelasi, , uji chi<sup>2</sup>, Uji t, dan Uji anova dengan jujur, kreatif, inovatif, mandiri, bertanggung jawab, dan religius

Soft skill/Karakter: Berfikir kritis, cerdas, teliti, jujur, kreatif, inovatif, mandiri bertanggung jawab, dan religius

**Matriks Pembelajaran:**

Minggu	LO	Pengalaman Belajar	Materi/Pokok Bahasan	Metode Strategi Pembelajaran	Kriteria/Teknik Penilaian	Daftar Pustaka
1	2	3	4	5	6	7
I - II	Berpikir kritis, cerdas, dan teliti dalam menjelaskan konsep dasar dan ruang lingkup statistika	Mendengarkan, memperhatikan, mencatat dan mendiskusikan	1. Konsep dasar statistika 2. Ruang lingkup statistika 3. Metode statistika	Ceramah dan Diskusi	Sikap Lisan	1,4,5
III	Memiliki pengetahuan tentang klasifikasi dan distribusi data	Mendengarkan, memperhatikan, mencatat dan mendiskusikan	1. Klasifikasi data 2. Distribusi data	Ceramah, Diskusi, latihan	Sikap Lisan	1,2,3
IV	Mampu melakukan perhitungan ukuran tendensi sentral	Mendengarkan, memperhatikan, mereview dan	1. Rata-rata (mean) 2. Median 3. Modus	Ceramah, Diskusi, latihan	Sikap Kinerja	1 s/d 6

		mendiskusikan				
V	Memiliki kemampuan menentukan ukuran sebaran	Mendengarkan, melihat dan mendiskusikan	<ol style="list-style-type: none"> <li>4. Rentang</li> <li>5. Deviasi rata-rata</li> <li>6. Varian</li> <li>7. Deviasi standar</li> </ol>	Ceramah, Diskusi, latihan	Sikap Lisan Kinerja	1 s/d 6
VI	Memiliki pemahaman yang luas, cerdas dan teliti tentang hipotesis. Pengenalan Program SPSS	Mendengarkan, melihat dan mendiskusikan	<ol style="list-style-type: none"> <li>1. Pengertian hipotesis</li> <li>2. Jenis hipotesis</li> <li>3. SPSS</li> </ol>	Ceramah, Diskusi, latihan	Sikap Kinerja Tugas	1 s/d 6
VII- VIII	Memiliki pemahaman yang luas dan teliti tentang uji prasyarat (analisis normalitas),	Mendengarkan, melihat dan mendiskusikan	<ol style="list-style-type: none"> <li>1. Normalitas</li> <li>2. Homogen dan Linearitas <ol style="list-style-type: none"> <li>a. Pengertian</li> <li>b. Penggunaan</li> <li>c. Perhitungan</li> </ol> </li> </ol>	Ceramah, Diskusi, latihan	Sikap Kinerja Tugas	1 s/d 6

	homogenitas, dan linearitas, dan menerapkan menggunakan program SPSS.					
/IX	<b>Ujian Tengah Semester</b>					
X-XI	Memiliki kemampuan dalam mengolah data dengan uji korelasi	Mendengarkan, melihat dan mendiskusikan	1. Spearman 2. Pearson. 3. Parsial.	Ceramah,Diskusi, latihan	Sikap Tulisan Kinerja	2 dan 4
XII	Memiliki kemampuan tentang uji chi <sup>2</sup>	Mendengarkan, melihat, dan mendiskusikan	1. Uji chi <sup>2</sup>	Ceramah,Diskusi, latihan	Sikap Tulisan Kinerja	1 dan 4
XIII - XIV	Memiliki kemampuan tentang uji t	Mendengarkan, melihat, dan mendiskusikan	1. Uji t satu sampel bebas 2. Uji t dua sampel berpasangan 3. Uji t dua sampel bebas	Ceramah,Diskusi, latihan	Sikap Tulisan Kinerja	1 s/d 6

XV – XVI	Memiliki kemampuan tentang uji Anova	Mendengarkan, melihat, dan mendiskusikan	1. Anova one way	Ceramah, Diskusi, latihan	Sikap Tulisan Kinerja	1 s/d 6
XVII	<b>Ujian Akhir Semester</b>					



### **Daftar Pustaka**

1. Sugiono.2010. Metode Penelitian Kuantitatif, kualitatif dan R & D.
2. Suharsimi Arikunto.1987. Prosedur Penelitian Suatu Pendekatan Praktis.
3. Nasir. 1983. Metode Penelitian. Ghalia Indonesia. Jakarta.
4. Singgih Santoso. 2002. SPSS Statistik Multivariat. Elek Media Komputindo. Jakarta.
5. Syamsul Ma'arif dan Hendri Tanjung. 2003. Teknik-Teknik kuantitatif untuk manajemen. Grasindo. Jakarta.

# MATERI KULIAH: STATISTIKA

**Bahan Kajian** : Konsep Dasar dan Ruang Lingkup Statistika  
**Program Studi** : Pendidikan Teknik Informatika  
**Minggu ke-** : 1 dan 2

## A. Pembelajaran :

### 1. Learning Outcomes

Berpikir kritis, cerdas, dan teliti dalam menjelaskan konsep dasar dan ruang lingkup statistika

### 2. Materi

- a) Konsep dasar statistika
- b) Ruang lingkup statistika
- c) Metode statistika

### 3. Uraian Materi

#### A. Konsep Dasar Statistika

Hampir dalam tiap bidang baik pemerintahan, pendidikan, perekonomian, perindustrian, atau lainnya akan menghadapi persoalan yang diantaranya dinyatakan dengan angka-angka. Kumpulan angka-angka ini biasanya disusun dalam tabel atau daftar disertai diagram atau grafik. Kumpulan angka-angka mengenai suatu masalah yang dapat memberi gambaran mengenai masalah tersebut dinamakan statistik, seperti statistik penduduk, statistik kelahiran, statistik pendidikan dan lain-lain. Statistik juga diartikan sebagai ukuran yang dihitung dari sekumpulan data dan merupakan wakil dari data itu.

Statistik berasal dari bahasa Latin, yaitu status yang berarti negara dan digunakan untuk urusan negara. Hal ini dikarenakan pada mulanya, statistik hanya digunakan untuk menggambarkan keadaan dan menyelesaikan masalah yang berhubungan dengan kenegaraan saja seperti : perhitungan banyaknya penduduk, pembayaran pajak, gaji pegawai, dan lain sebagainya. Seiring dengan perkembangan zaman, statistik mulai mencakup hal-hal yang lebih

luas. Sekarang ruang lingkup statistik tidak hanya bertumpu pada angka-angka untuk pemerintahan saja, tetapi telah mengambil bagian di berbagai bidang kehidupan, termasuk kegiatan berbagai bidang penelitian, seperti teknik, kedokteran, pendidikan dan psikologi, pertanian, sosial, dan sains. Berikut ini beberapa pengertian statistik sesuai dengan perkembangannya. Pengertian pertama: statistik adalah sekumpulan angka untuk menerangkan sesuatu, baik angka yang belum tersusun (masih acak) maupun angka-angka yang sudah tersusun dalam suatu daftar atau grafik. Berdasarkan pengertian ini, statistik diartikan dalam arti sempit, yaitu keterangan ringkas berbentuk angka-angka. Contoh: statistik penduduk, yang berarti keterangan mengenai penduduk berupa angka-angka dalam bentuk ringkas, seperti jumlah penduduk dan rata-rata umur penduduk.

Pengertian kedua: statistik adalah sekumpulan cara dan aturan tentang pengumpulan, pengolahan, analisis, serta penafsiran data yang terdiri dari angka-angka. Selanjutnya pengertian ketiga, statistik adalah sekumpulan angka yang menjelaskan sifat-sifat data atau hasil pengamatan. Berdasarkan pengertian kedua dan ketiga ini, statistik sudah diartikan dalam arti yang luas dan sudah merupakan suatu metode atau ilmu, yaitu metode atau ilmu yang mempelajari cara pengumpulan, pengolahan, penganalisisan, penafsiran, dan penarikan kesimpulan dari data yang ada. Contoh: Seorang pemilik pabrik roti merek SELERA ingin mengetahui jumlah bungkus roti tersebut yang dikonsumsi tiap rumah tangga per bulan, di sebuah kelurahan. Di kelurahan tersebut tinggal 1.000 rumah tangga. Dari 1.000 rumah tangga tersebut dipilih 100 rumah tangga sebagai sampel. Selanjutnya dari 100 sampel itu, data dikumpulkan, diolah, dan dianalisis. Akhirnya diketahui bahwa rata-rata jumlah bungkus yang digunakan tiap rumah tangga setiap bulannya berkisar 10 sampai 15 buah. Statistik dalam arti sempit mendeskripsikan atau menggambarkan mengenai data yang disajikan dalam bentuk (1) Tabel dan diagram, (2) Pengukuran tendensi sentral (rata-rata hitung, rata-rata ukur, dan rata-rata harmonik), (3) Pengukuran penempatan (median, kuartil, desil, dan presentil), (4) Pengukuran penyimpangan (range, rentangan antar kuartil,

rentangan semi antar kuartil, simpangan rata-rata, simpangan baku, variansi, koefisien variansi dan angka baku), dan (5) Angka indeks.

Statistik dalam arti luas adalah suatu alat untuk mengumpulkan data, mengolah data, menarik kesimpulan, membuat tindakan berdasarkan analisis data yang dikumpulkan atau statistika yang digunakan menganalisis data sampel dan hasilnya dimanfaatkan untuk generalisasi pada populasi. Selanjutnya, untuk memperjelas pengertian tersebut di atas, beberapa pengertian yang dikemukakan oleh beberapa ahli, antara lain: (1) Statistik digunakan untuk membatasi cara-cara ilmiah untuk mengumpulkan, menyusun, meringkas, dan menyajikan data penyelidikan. Lebih jauh dinyatakan bahwa statistik merupakan cara untuk mengolah data dan menarik kesimpulan-kesimpulan yang teliti dan keputusan-keputusan yang logis dari pengolahan data tersebut (Sutrisno Hadi, 1987), (2) Statistik adalah pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan penganalisisannya, dan penarikan kesimpulan berdasarkan kumpulan data dan analisis yang dilakukan. Statistik adalah metode yang memberikan cara-cara guna menilai ketidaktentuan dan penarikan kesimpulan yang bersifat induktif. Berdasarkan pengertian-pengertian mengenai statistik, terlihat adanya pergeseran pengertian, dari pengertian yang sempit ke pengertian yang luas. Oleh karena itu, pengertian statistik yang lebih jelas dan melingkupi pengertian, baik yang sempit maupun yang luas berikut ini. Statistik adalah ilmu yang mempelajari tentang seluk-beluk data, yaitu tentang pengumpulan, pengolahan/analisis, penafsiran, dan penarikan kesimpulan dari data yang berbentuk angka-angka.

## **B. Ruang Lingkup Statistika**

Dalam kehidupan yang modern sekarang ini, dengan ciri utama adalah globalisasi, statistik tidak diragukan lagi peranannya dalam membantu memudahkan kehidupan manusia. Lebih jelasnya, peranan statistik antara lain terlihat dalam kehidupan sehari-hari, dalam kegiatan ilmiah, dan kegiatan proses belajar mengajar, dan dalam kegiatan ilmu pengetahuan.

1. Dalam kehidupan sehari-hari, statistik memiliki peranan sebagai penyedia bahan-bahan atau keterangan-keterangan berbagai hal untuk diolah dan ditafsirkan. Contoh: angka kenaikan jumlah konsumsi pangan, tingkat biaya hidup, tingkat import terigu, dan tingkat pendapatan.
2. Dalam penelitian ilmiah, statistik memiliki peranan sebagai penyedia data untuk mengemukakan atau menemukan kembali keterangan-keterangan yang seolah-olah tersembunyi dalam angka-angka statistik.
3. Dalam kegiatan proses belajar mengajar, statistik banyak membantu dalam menganalisis soal-soal yang diberikan dalam kegiatan pembelajaran. Contoh: perbandingan banyaknya siswa perempuan dan laki-laki di kelas I, rerata prestasi siswa tata boga di kelas XI, dan besarnya indeks objektivitas sekolah SMK' dalam mengikuti Ujian Nasional Matematika.
4. Dalam kegiatan ilmu pengetahuan, statistik memiliki peranan sebagai sarana analisis dan interpretasi dari data kuantitatif ilmu pengetahuan, sehingga diperoleh suatu kesimpulan dari berbagai data tersebut. Semakin pentingnya peranan statistik pada berbagai bidang dalam kehidupan modern, menimbulkan berbagai macam cabang ilmu baru yang merupakan gabungan antara ilmu tersebut dengan statistik atau penerapan statistik dalam ilmu tersebut. Cabang-cabang ilmu baru tersebut, antara lain: (1) ekonometrika, merupakan gabungan antara ilmu ekonomi dengan statistik; (2) sosiometri, merupakan gabungan antara ilmu sosiologi dengan statistik; dan (3) psikometri, merupakan gabungan antara ilmu psikologi dengan statistik.

Statistik perlu diketahui dan dipelajari karena statistik berperan sebagai alat bantu dalam hal-hal berikut ini.

1. Menjelaskan hubungan antara variabel-variabel  
Variabel atau peubah merupakan sesuatu yang nilainya bervariasi (tidak tetap), seperti harga, produksi, hasil penjualan, umur, dan tinggi. Dengan menggunakan statistik, variabel-variabel tersebut dapat dijelaskan hubungannya. Misalnya, hubungan antara hasil tes seleksi dengan indeks

prestasi siswa, kecepatan membaca dengan ketelitian menghitung. Analisis korelasi dan regresi mampu memberikan jawaban yang terbaik.

2. Membuat rencana dan ramalan

Rencana dan ramalan merupakan dua hal yang diperlukan dalam pelaksanaan sesuatu, sehingga dapat diperoleh hasil yang baik dan berkualitas. Oleh karena itu, rencana dan ramalan harus baik pula. Dengan statistik, rencana dan ramalan dapat dibuat sebaik mungkin.

3. Mengatasi berbagai perubahan

Perubahan-perubahan yang terjadi dalam suatu pengambilan keputusan, tidak mungkin dapat diabaikan atau dihindarkan, supaya pihak-pihak lain tidak ada yang dirugikan. Dengan statistik, perubahan-perubahan yang mungkin terjadi dapat diantisipasi sedini mungkin. Sebagai contoh, ketua Serikat Pekerja ingin mengadakan perjanjian dengan pimpinan sebuah perusahaan. Agar upah riil tidak mengalami perubahan dan buruh tidak dirugikan maka ketua serikat pekerja perlu memperhatikan perkembangan indeks harga yang menyangkut perubahan seluruh harga barang untuk periode saat itu dari periode sebelumnya. Perhitungan angka indeks dapat memberikan jawabannya.

4. Membuat keputusan yang lebih baik

Keputusan yang baik dan rasional amat diperlukan dalam menjaga kelancaran sebuah aktivitas kerja supaya kelestarian dari sebuah usaha dapat terjamin. Dengan statistik, keputusan yang baik dan rasional dapat dihasilkan. Sebagai contoh, seorang kepala sekolah dihadapkan pada kondisi yang tidak menentu dari prestasi para siswanya. Kepala sekolah harus dapat mengambil sikap atau tindakan tertentu, misalnya melihat grafik perkembangan siswanya, memotivasi para guru untuk bekerja lebih giat, memperbaiki kualitas soal ujian berdasarkan analisis validitas butir, dan lain sebagainya yang terfokus pada analisis data. Teori keputusan dan uji hipotesis dapat membantu pelaksanaannya.

Statistik mempunyai fungsi, antara lain sebagai:

1. Bank data untuk menyediakan data untuk diolah dan diinterpretasikan agar dapat digunakan untuk menerangkan keadaan yang perlu diketahui atau diungkap.
2. Alat *quality control* untuk membantu standardisasi dan sekaligus sebagai alat pengawasan.
3. Alat analisis, merupakan suatu metode penganalisisan data.
4. Pemecahan masalah dan pembuatan keputusan, sebagai dasar penetapan kebijakan dan langkah lebih lanjut untuk mempertahankan, mengembangkan perusahaan dalam perolehan keuntungan.

Contoh soal : jelaskan peranan Statistika dalam pengembangan bidang ilmu lainnya serta berikan contoh penggunaannya!

Jawab:

Peranan Statistika dalam pengembangan bidang Ilmu lainnya adalah sebagai:

- a. Alat untuk menghitung besarnya anggota sampel yang diambil dari suatu populasi, sehingga jumlah sampel yang dibutuhkan akan lebih dapat dipertanggungjawabkan
- b. Alat untuk menguji validitas dan reliabilitas instrumen sebelum instrumen tersebut digunakan dalam penelitian
- c. Sebagai teknik untuk menyajikan data, sehingga data lebih komunikatif, misalnya melalui tabel, grafik, atau diagram
- d. Alat untuk menganalisis data seperti menguji hipotesis yang diajukan dalam penelitian.

Adapun beberapa contoh penggunaan Statistika dalam bidang Ilmu lainnya antara lain:

- a. Dalam ilmu ekonomi, statistic berperan sebagai alat analisa bursa efek saham, kurs rupiah ataupun terhadap kondisi pendapatan dan pengeluaran dalam suatu kelompok, penduduk atau masyarakat.

- b. Dalam ilmu kedokteran, statistika berperan sebagai alat analisa terhadap suatu gejala yang akan dijadikan dasar penyimpulan jenis penyakit.
- c. Dalam ilmu geofisika, statistika berperan sebagai landasan meramalkan cuaca dengan menggunakan gejala atau kondisi yang ada.
- d. Dalam ilmu biologi, statistika digunakan untuk melakukan pengelompokan / tingkatan makhluk hidup.
- e. Dalam ilmu psikologi, para psikolog banyak menggunakan statistika untuk membaca hasil pengamatan baik melalui tes maupun obserbasi lapangan.

Contoh lain dari peran Statistika yakni:

- a. Pemerintah menggunakan statistika untuk menilai hasil pembangunan masa lalu dan merencanakan masa mendatang
- b. Pimpinan menggunakannya untuk pengangkatan pegawai baru, pembelian peralatan baru, peningkatan kemampuan karyawan, perubahan sistem kepegawaian, dsb.
- c. Para pendidik sering menggunakannya untuk melihat kedudukan siswa, prestasi belajar, efektivitas metoda pembelajaran, atau media pembelajaran.

Tujuan pembelajaran Statistika adalah sebagai suatu pendekatan modern untuk menyajikan mengenai konsep-konsep dasar dan metode statistik secara lebih jelas dan langsung dapat membantu seseorang didalam pengembangan daya kritik dalam suatu kegiatan pengambilan keputusan dengan menggunakan cara-cara kuantitatif, dalam rangka membantu memberi bobot didalam mengambil keputusan. Sedangkan kegunaan Statistik sebagai ilmu pengetahuan adalah:

- a. Untuk menggambarkan keadaan, baik secara umum maupun secara khusus;
- b. Untuk memperoleh gambaran tentang perkembangan (pasang-surut) dari waktu ke waktu;
- c. Untuk mengetahui perbandingan (membandingkan) antara gejala yang satu dengan gejala yang lain;



- d. Untuk menilai keadaan dengan jalan menguji perbedaan antara gejala yang satu dengan gejala yang lain;
- e. Untuk menilai keadaan dengan jalan mencari hubungan antara gejala yang satu dengan gejala yang lain;
- f. Untuk menjadi dasar atau pedoman, baik di dalam menarik kesimpulan, mengambil keputusan, serta memperkirakan terjadinya sesuatu hal atas dasar bahan-bahan keterangan (data) yang telah berhasil dihimpun, dan lain sebagainya.

### **C. Metode Statistik**

#### **1. Jenis statistik berdasar cara pengolahan data dibagi atas dua bagian.**

##### **a. Statistik deskriptif**

Statistik deskriptif adalah bagian dari statistik yang mempelajari cara pengumpulan dan penyajian data sehingga mudah dipahami. Statistik deskriptif hanya berhubungan dengan hal menguraikan atau memberikan keterangan-keterangan mengenai suatu data atau keadaan atau fenomena. Dengan kata lain, statistik deskriptif hanya berfungsi menerangkan keadaan, gejala, atau persoalan. Beberapa contoh pernyataan yang termasuk dalam ruang lingkup statistik deskriptif, antara lain:

- 1) Sekurang-kurangnya 5 % dari kebakaran yang terjadi di kota “Padang“, yang dilaporkan tahun lalu diakibatkan oleh tindakan-tindakan sengaja yang tidak bertanggung jawab.
- 2) Sebanyak 20 % di antara semua pasien yang menerima suntikan obat tertentu, ternyata kemudian menderita efek samping obat tersebut.
- 3) Penarikan kesimpulan pada statistik deskriptif (apabila ada) hanya ditujukan pada kumpulan data yang ada, didasarkan atas ruang lingkup bahasannya.

Berdasarkan atas ruang lingkup bahasannya, statistik deskriptif mencakup hal berikut ini.

- 1) Distribusi frekuensi beserta bagian-bagiannya, seperti grafik distribusi (histogram, poligon frekuensi, dan Ogive); b) ukuran nilai pusat (rata-rata,

median, modus, kuartil, dan sebagainya); c) ukuran dispersi (jangkauan, simpangan rata-rata, variasi, simpangan baku, dan sebagainya); d) kemencengan dan keruncingan kurva.

- 2) Angka indeks
- 3) *Time series*/deret waktu atau data berkala
- 4) Korelasi dan regresi sederhana.

#### b. Statistik inferensial

Statistik inferensial adalah serangkaian teknik yang digunakan untuk mengkaji, menaksir dan mengambil kesimpulan sebgaiian data (data sampel) yang dipilih secara acak dari seluruh data yang menjadi subyek kajian (populasi). Statistik inferensial berhubungan dengan pendugaan populasi dan pengujian hipotesis dari suatu data atau keadaan atau fenomena. Dengan kata lain, statistik inferensial berfungsi meramalkan dan mengontrol keadaan atau kejadian. Berikut ini beberapa contoh pernyataan yang termasuk dalam cakupan statistik inferensial.

- 1) Akibat penurunan produksi minyak oleh negara-negara penghasil minyak dunia, diramalkan harga minyak akan menjadi dua kali lipat pada tahun-tahun mendatang.
- 2) Dengan mengasumsikan bahwa keracunan makanan dari pengolahan makanan menggunakan jasa katering “A” kurang dari 30 % akibat hygiene dan sanitasi yang rendah, maka penghasilan katering tersebut di akhir tahun nanti tidak akan lebih dari 1.500 rupiah per satu kotaknya.

Penarikan kesimpulan pada statistik inferensial ini merupakan generalisasi dari suatu populasi berdasarkan data (sampel) yang ada. Berdasar atas ruang lingkup bahasannya, maka statistik inferensial mencakup hal-hal berikut ini.

- 1) Probabilitas atau teori kemungkinan
- 2) Distribusi teoretis
- 3) Sampling dan sampling distribusi
- 4) Pendugaan populasi atau teori populasi
- 5) Uji hipotesis rerata
- 6) Analisis korelasi dan uji signifikansi

7) Analisis regresi untuk peramalan

8) Analisis varians; dan

9) Analisis kovarians

**2. Jenis statistik berdasar atas ruang lingkup penggunaan atau disiplin ilmu yang menggunakannya, statistik dapat dibagi atas beberapa macam.**

- a. Statistik pendidikan adalah statistik yang diterapkan atau digunakan dalam bidang ilmu pendidikan.
- b. Statistik perusahaan adalah statistik yang diterapkan atau digunakan dalam bidang perusahaan.
- c. Statistik ekonomi adalah statistik yang diterapkan atau digunakan dalam bidang ilmu ekonomi
- d. Statistik pertanian adalah statistik yang digunakan atau diterapkan dalam bidang ilmu pertanian.
- e. Statistik kesehatan adalah statistik yang digunakan atau diterapkan dalam bidang ilmu kesehatan.
- f. Statistik sosial adalah statistik yang diterapkan atau digunakan dalam bidang ilmu sosial.

**3. Jenis statistik berdasarkan bentuk parameternya (data yang sebenarnya), statistik dapat dibagi atas dua bagian.**

a. Statistik parametrik

Statistik parametrik adalah bagian statistik yang parameter populasinya mengikuti suatu distribusi tertentu, seperti distribusi normal dan memiliki varians yang homogen.

b. Statistik nonparametrik

Statistik nonparametrik adalah bagian statistik yang parameter populasinya tidak mengikuti suatu distribusi tertentu atau memiliki distribusi yang bebas dari persyaratan, dan variansnya tidak perlu homogen.

# MATERI KULIAH: STATISTIKA

**Bahan Kajian** : **Klasifikasi dan Distribusi Data**  
**Program Studi** : **Pendidikan Teknik Informatika**  
**Minggu ke-** : **3**

## A. Pembelajaran :

### 1. Learning Outcomes

Memiliki pengetahuan tentang klasifikasi dan distribusi data

### 2. Materi

- a) Klasifikasi data
- b) Distribusi data

### 3. Uraian Materi

#### A. Data

Data adalah bentuk jamak dari *datum*. Data merupakan kumpulan fakta atau angka atau segala sesuatu yang dapat dipercaya kebenarannya sehingga dapat digunakan sebagai dasar menarik suatu kesimpulan. Tidak semua angka dapat disebut data statistik. Angka dapat disebut data statistik apabila dapat menunjukkan suatu ciri dari suatu penelitian yang bersifat agregatif, serta mencerminkan suatu kegiatan lapangan tertentu. Data penelitian merupakan faktor yang sangat mempengaruhi teknik atau jenis penelitian yang akan digunakan.

Penggolongan data statistik dapat ditinjau dari :

- a) Variabel yang diteliti (segi sifat angkanya), data statistik dapat dibedakan menjadi dua golongan, yaitu data kontinyu dan data diskrit. Data kontinyu adalah data statistik yang angka-angkanya merupakan deretan angka yang sambung-menyambung.  
Data diskrit ialah data statistik yang tidak mungkin berbentuk pecahan.
- b) Cara menyusun angka, data statistik dapat dibedakan menjadi data nominal, data ordinal, dan data interval. Data nominal ialah data statistik yang cara menyusun angkanya didasarkan atas penggolongan atau klasifikasi tertentu. Data ordinal juga sering disebut dengan data urutan, yaitu data statistik yang cara menyusun angkanya didasarkan atas urutan kedudukan (ranking). Data interval ialah data statistik yang

terdapat jarak sama di antara hal-hal yang sedang diselidiki atau dipersoalkan.

- c) Bentuk angka, data statistik dapat dibedakan menjadi 2 (dua) macam, yaitu data tunggal (*un grouped data*) dan data kelompok atau data bergolong (*grouped data*).
- d) Sumber mana data tersebut diperoleh, data statistik dapat dibedakan menjadi dua macam, yaitu: data primer dan data sekunder. Data primer adalah data statistik yang diperoleh atau bersumber dari tangan pertama (*first hand data*). Sedangkan data sekunder adalah data statistik yang diperoleh dari tangan kedua (*second hand data*).

#### Data Menurut Waktu Pengumpulan

##### a. Data Cross Section

Data yang dikumpulkan dalam suatu periode waktu tertentu, biasanya menggambarkan keadaan atau kegiatan pada periode tersebut. (Data yang tidak bergantung pada waktu) Contoh: Hasil sensus penduduk pada tahun 1990 menggambarkan karakteristik penduduk Indonesia pada tahun 1990 (Jenis Kelamin-Tingkat Ekonomi)

##### b. Data Time Series

Data yang dikumpulkan dari waktu-kewaktu dengan tujuan untuk mengetahui perkembangan suatu kegiatan dari waktu-kewaktu. (Data bergantung pada waktu) Contoh: Hasil produksi produk nugget dari tahun ke tahun

#### Syarat data yang baik

Syarat data yang baik adalah (a) Data harus objektif (sesuai dengan keadaan sebenarnya), (b) Data harus representative (data harus mewakili objek yang diamati), (c) Data harus *up to date*, dan (d) Data harus relevan dengan masalah yang akan dipecahkan, (e) Kesalahan baku kecil (tingkat ketelitian).

## **B. Variabel**

Variabel adalah sebuah simbol, yang dapat menyandang setiap nilai dari suatu himpunan nilai yang disebut sebagai domain dari variabel tersebut. Secara umum, variabel dibagi atas 2 (dua) jenis, yaitu variabel kontinu (*continous variabel*) dan variabel deskrit (*descrete variabel*). Variabel kontinu merupakan variabel yang dapat ditentukan nilainya dalam jarak jangkauan tertentu dengan desimal yang tidak terbatas. Contoh variabel kontinu yaitu tinggi seseorang yang dapat bernilai 62 cm, 67,5 cm atau 68,45678 cm, bergantung pada tingkat akurasi pengukurannya. Data yang dijelaskan melalui variabel kontinu disebut data kontinu

Variabel diskrit adalah konsep yang nilainya tidak dapat dinyatakan dalam bentuk pecahan atau desimal di belakang koma. Contoh variabel diskrit, sejumlah  $n$  anak dalam sebuah keluarga, yang bernilai bisa salah satu dari 0, 1, 2, 3, ... tetapi tidak mungkin 2,5 atau 3,4567. Data yang dijelaskan melalui variabel diskrit disebut data diskrit.

Variabel dapat juga dibagi sebagai variabel dependen dan variabel bebas. Apabila ada hubungan antara dua variabel, misalnya antara variabel  $Y$  dan variabel  $X$ , dan jika variabel  $Y$  disebabkan oleh variabel  $X$ , maka variabel  $Y$  adalah variabel dependen dan variabel  $X$  adalah variabel bebas.

Variabel dapat dilihat sebagai variabel aktif dan variabel atribut. Variabel aktif adalah variabel yang dimanipulasikan oleh peneliti. Variabel atribut merupakan variabel-variabel yang tidak dapat dimanipulasikan atau sukar dimanipulasi. Variabel-variabel atribut umumnya merupakan karakteristik manusia seperti; inteligensia, jenis kelamin, status sosial, pendidikan, sikap, dan sebagainya.

## **C. Skala Pengukuran**

Skala merupakan hasil pengukuran yang terdiri atas beberapa jenis skala yang bervariasi. Pengukuran adalah pemberian angka terhadap objek atau fenomena menurut aturan tertentu. Tiga buah kata kunci yang diperlukan dalam memberikan definisi terhadap konsep pengukuran. Kata-kata kunci tersebut adalah angka, penetapan, dan aturan. Pengukuran yang baik, harus mempunyai

sifat *isomorphism* dengan realita. Prinsip *isomorphism*, artinya terdapat kesamaan yang dekat antara realitas sosial yang diteliti dengan "nilai" yang diperoleh dari pengukuran. Oleh karena itu, suatu instrumen pengukur dipandang baik apabila hasilnya dapat merefleksikan secara tepat realitas dari fenomena yang hendak diukur.

Ada empat skala pengukuran data, yaitu: nominal, ordinal, interval, dan rasio.

- a) Ukuran nominal, adalah ukuran yang paling sederhana, dimana angka yang diberikan kepada objek mempunyai arti sebagai label saja, dan tidak menunjukkan tingkatan apa-apa. Ukuran nominal adalah data yang dinyatakan dalam bentuk kategori. Misalnya bentuk bank syariah di Indonesia: Bank Umum Syariah diberi kategori 2; BPR Syariah diberi kategori 1. Laki-laki diberi kode 1, dan wanita diberi kode 0. Ukuran nominal hanya memberikan suatu informasi yang bersifat dasar, kategorial, dan mentah sehingga tidak dapat dinotasikan dalam fungsi matematika. Contoh lain ukuran nominal adalah jenis pekerjaan (nelayan, petani, PNS), agama (Islam, Nasrani, Hindu, Budha), asal (Padang, Medan, Jakarta, Makasar) dan lain-lain.
- i. Ukuran ordinal adalah angka yang diberikan mengandung pengertian tingkatan. Ukuran nominal digunakan untuk mengurutkan objek dari yang terendah ke yang tertinggi atau sebaliknya.

Contoh ukuran ordinal:

Urutkan merk sepeda motor berikut dari yang paling anda sukai.

Merk	Ranking
Yamaha	.....
Honda	.....
Suzuki	.....
Kawasaki	.....

Contoh lain yaitu huruf mutu (A, B, C, D, E), Tingkat warna (putih, merah muda, merah, merah tua).

- ii. Ukuran interval adalah mengurutkan orang atau objek berdasarkan suatu atribut. Selain itu, juga memberikan informasi tentang interval antara satu

orang atau objek dengan orang atau objek lainnya. Interval atau jarak yang sama pada skala interval dipandang sebagai mewakili interval atau jarak yang sama pula pada objek yang diukur.

Contoh ukuran interval:

Beri penilaian (rate) empat merk sepeda motor ini berdasarkan 5 skala penilaian : 1 (sangat jelek) s/d 5 (sangat bagus).

Merk	Nilai
Yamaha	.....
Honda	.....
Suzuki	.....
Kawasaki	.....

- iii. Ukuran rasio, adalah ukuran yang mencakup semua ukuran sebelumnya ditambah dengan satu sifat lain, yaitu ukuran ini memberikan keterangan tentang nilai absolut dari objek yang diukur. Ukuran rasio mempunyai titik nol, karena itu interval jarak tidak dinyatakan dengan beda angka rata-rata satu kelompok dibandingkan dengan titik nol. Karena ada titik nol tersebut, maka ukuran rasio dapat dibuat perkalian ataupun pembagian. Angka pada skala rasio menunjukkan nilai sebenarnya dari objek yang diukur.

Contoh skala rasio adalah usia (1, 12, 34, 40), jumlah penjualan (20, 100, 21, 82, 76), penghasilan, laba, gaji, skor ujian dan sebagainya.

Perbedaan keempat skala pengukuran tersebut dapat disajikan pada Tabel 1 sebagai berikut :

**Tabel 1.1. Perbedaan Skala Pengukuran**

Skala pengukuran				
	Pengkelasan	Pengurut	Pembeda	Pembanding
Nominal				
Ordinal				
Interval				
Rasio				

Jenis data atau ukuran yang digunakan dalam penelitian sangat terkait erat dengan teknik analisis yang digunakan, karena pada statistik tidak semua data bisa digunakan. Misalnya untuk data berjenis ordinal dan nominal maka teknik



analisis yang digunakan adalah statistik non parametris, sedangkan pada data interval dan rasio, teknik yang digunakan adalah statistik parametris.

#### **D. Distribusi Data**

##### **1. PENGERTIAN**

Data kuantitatif yang dikumpulkan dari lapangan (data mentah), nilainya tidak selalu sama atau seragam tetapi bervariasi dari satu pengamatan ke pengamatan yang lain, misalnya data hasil produksi, data hasil penjualan, data tingkat konsumsi dan lain-lain. Jika data hasil pengamatan di lapangan mempunyai jumlah yang besar maka data mentah tersebut perlu diolah dengan cara meringkas data tersebut dan didistribusikan ke dalam kelas atau kategori. Suatu tabel yang berisi susunan data yang terbagi ke dalam beberapa frekuensi kelas disebut Distribusi Frekuensi atau Tabel Frekuensi. Dengan disajikannya data dalam bentuk distribusi frekuensi maka akan memudahkan bagi pihak yang berkepentingan terhadap data tersebut untuk melakukan analisis data, dibandingkan jika data yang disajikan masih berupa data mentah dan dalam jumlah yang banyak.

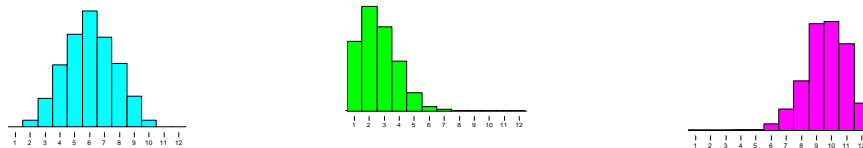
Distribusi frekuensi adalah yang merupakan penyusunan data ke dalam kelas-kelas tertentu dimana setiap individu/item hanya termasuk kedalam salah satu kelas tertentu saja (pengelompokkan data berdasarkan kemiripan ciri). Distribusi frekwensi dapat diartikan sebagai pengelompokkan data kedalam kategori-kategori atau kelas-kelas dengan interval tertentu. Banyaknya data yang berada dalam kelas masing-masing disebut frekwensi. Distribusi frekwensi ini dimaksudkan untuk lebih mudah menganalisis data yang dimaksud atau dengan tujuan untuk mengatur data mentah (belum dikelompokkan) ke dalam bentuk yang rapi tanpa mengurangi inti informasi yang ada.

Distribusi Frekuensi merupakan suatu tabel yang menunjukkan frekuensi kemunculan data atau frekuensi relatifnya yang berguna untuk meringkas data numerik maupun kategori.

- Untuk data diskret atau data kategori, banyaknya nilai yang dihitung kemunculannya biasanya sesuai dengan banyaknya nilai data yang berbeda dari data diskret atau kategori tersebut
- Untuk data kontinu, biasanya dibuat kelas interval 5-20 banyaknya.

Bentuk distribusi standar ada tiga (Gambar 1) yaitu :

- Simetris : jika penyebaran data sebelah kiri dan kanan dari nilai rata-rata populasi adalah sama.
- Menjulang ke kanan : jika data mengumpul dinilai-nilai yang kecil (disebelah kiri) dan sisanya (data dengan nilai-nilai besar) menyebar di sebelah kanan.
- Menjulang ke kiri : jika data mengumpul dinilai-nilai yang besar (disebelah kanan) dan sisanya (data dengan nilai-nilai kecil) menyebar di sebelah kiri.



*Gambar 1: Bentuk Distribusi Standar*

Distribusi frekuensi merupakan tabel ringkasan data yang menunjukkan frekuensi/ banyaknya item/obyek pada setiap kelas yang ada. Tujuan distribusi frekuensi adalah untuk mendapatkan informasi lebih dalam tentang data yang ada yang tidak dapat secara cepat diperoleh dengan melihat data aslinya.

## 2. ISTILAH-ISTILAH DALAM DISTRIBUSI FREKUENSI.

Istilah yang sering digunakan dalam distribusi frekuensi yaitu :

- 1) Kelas adalah penggolongan data yang dibatasi dengan nilai terendah dan nilai tertinggi yang masing-masing dinamakan batas kelas. Batas Kelas adalah nilai batas dari pada tiap kelas dalam sebuah distribusi, terbagi menjadi *States class limit* dan *Class Bounderies* (Tepi kelas). *Stated Class Limit* adalah batas-batas kelas yang tertulis dalam distribusi frekuensi, terdiri dari *Lower Class Limit* (Batas bawah kelas) dan *Upper Class Limit* (Batas atas kelas. *Class*

*Bounderies* (Tepi kelas) adalah batas kelas yang sebenarnya, terdiri dari *Lower class boundary* (batas bawah kelas yang sebenarnya) dan *upper class boundary* (batas atas kelas yang sebenarnya).

- 2) *Class Interval* atau Panjang Kelas atau Lebar kelas merupakan lebar dari sebuah kelas dan dihitung dari perbedaan antara kedua tepi kelasnya.
- 3) *Mid point* atau *Class Mark* atau Titik tengah merupakan rata-rata hitung dari kedua batas kelasnya atau tepi kelasnya.

### 3. JENIS DISTRIBUSI FREKWENSI

Distribusi frekwensi dapat dibedakan menjadi :

1. Distribusi frekwensi tunggal yaitu jika kategori atau kelas untuk mendistribusikan data bernilai tunggal. Contoh :

Data produksi susu dari seekor sapi (liter) : 0, 16, 4, 5, 15, 0, 11, 6, 12, 7, 12, 34, 19, 18, 17, 21, 10, 27, 6, 10, 10, 23, 17, 27, 6, 16, 21, 11, 12, 22, 16. Distribusi frekwensi tunggalnya :

*Tabel 1. Frekuensi Tunggal*

No.	Nilai	Frekwensi	Frekwensi kumulatif
1	0	2	2
2	4	1	3
3	5	1	4
4	6	3	7
5	7	1	8
6	10	3	11
7	11	2	13
8	12	3	16
9	15	1	17
10	16	3	20
11	17	2	22
12	18	1	23
13	19	1	24
14	21	2	26
15	22	1	27
16	23	1	28
17	27	2	30
18	34	1	31
Jumlah		31	

2. Distribusi frekwensi data yang dikelompokkan ( interval ).

Distribusi frekwensi data yang dikelompokkan ini digunakan untuk data yang banyak sekali dengan nilai-nilai yang sangat variatif. Karena distribusi frekwensi data yang dikelompokkan ini paling sering digunakan, maka istilah distribusi frekwensi sering ditujukan untuk distribusi frekwensi data yang dikelompokkan.

Pengelompokan nilai-nilainya sangat tergantung pada **kondisi data serta keperluannya**, penyusunan tabel distribusi frekuensi dapat dilakukan dengan beberapa tahap dan secara umum ada teori yang telah diturunkan oleh H.A. Sturges, langkah-langkah penyusunan tabel distribusi sebagai berikut :

- ) Mengurutkan data (bila diperlukan)
- ) Menentukan jumlah kelas : untuk menentukan jumlah kelas dapat digunakan

$$k = 1 + 3.322 \log n \dots\dots\dots 1$$

k = banyaknya kelas.

n = banyaknya data.

Contoh : lihat data produksi susu diatas.

n = banyaknya data = 31, maka banyaknya kelas :

$$k = 1 + 3.322 \log 31 = 1 + 4.95 = 1 + 5 = 6$$

- ) Menentukan jarak kelas (range) : range adalah jarak data terkecil sampai data terbesar atau selisih data terbesar dengan data terkecil

:

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

R = jarak kelas

Contoh : lihat data produksi susu diatas.

Nilai maksimum = 34, dan nilai minimum = 0,

maka :  $R = 34 - 0 = 34$

- ) Menentukan interval kelas : interval adalah panjang kelas yang nilainya diperoleh dari nilai range dibagi dengan nilai jumlah kelas

$$i = \frac{R}{k}$$

$i$  = interval kelas.

$R$  = range kelas ,

$k$  = jumlah kelas.

- ) Menentukan batas-batas kelas:  
Tepi bawah kelas = batas bawah kelas - 0,5 (skala terkecil)  
Tepi atas kelas = batas atas kelas + 0,5 (skala terkecil)  
Panjang interval kelas = Tepi atas kelas - tepi bawah kelas
- ) Menentukan titik tengahnya =  $\frac{1}{2}$  ( Batas atas kelas + batas bawah kelas)
- ) Memasukkan data ke dalam kelas-kelas yang sesuai dengan memakai sistem Tally atau Turus.
- ) Menyajikan distribusi frekuensi : isi kolom frekuensi sesuai dengan kolom Tally / Turus.
- ) Menentukan kelas ; dalam menentukan kelas yang harus diperhatikan adalah bahwa semua data harus dapat masuk dalam kelas tersebut dan tidak boleh terdapat data yang tersisa atau tidak dapat masuk dalam kelas yang telah ditentukan
- ) Mencari frekuensi masing-masing kelas : setelah data dapat masuk semua ke dalam kelas yang telah ditentukan maka langkah selanjutnya adalah menjumlahkan data masing-masing kelas atau disebut mencari frekuensi masing-masing kelas

Contoh : lihat data produksi susu diatas.

$R = 34$ ,

$k = 6$ , maka :

$i = 34 / 6 = 5,666$  .

Karena umumnya diambil harga bulat, maka  $i = 5$ , atau  $i = 6$

Umumnya untuk kelaziman interval kelas diambil nilai 5, atau kelipatannya.

Jika diambil interval kelas = 5, maka distribusi dari data diatas adalah sebagai berikut :

Distribusi frekwensi data yang dikelompokkan :

**Tabel 2. Frekwensi Data Yang Dikelompokkan.**

No	Kelas	Frekwensi	Frekwensi kumulatif
1	0 - 4	3	3
2	5 - 9	5	8
3	10 - 14	8	16
4	15 - 19	8	24
5	20 - 24	4	28
6	25 - 29	2	30
7	30 - 34	1	31
Jumlah		31	

Distribusi frekwensi lainnya berkaitan dengan teknik penyajiannya, seperti : distribusi frekwensi relatif, distribusi frekwensi fumulatif dan distribusi frekwensi relatif kumulatif.

3. Distribusi frekuensi relatif. Distribusi frekuensi relatif, yaitu suatu distribusi frekuensi yang fekuensi tiap kelas tidak dinyatakan dalam angka absolut, tetapi dalam angka relatif atau prosentase. Bentuk umum dari tabel distribusi frekuensi relatif :

Nilai Data	Frekuensi Relatif (%)
a - b	$g_1$
c - d	$g_2$
e - f	$g_3$
g - h	$g_4$
i - j	$g_5$
<b>Jumlah</b>	100

dengan frekuensi relatif kelas ke  $i$ ;  $g_i = (f_i/\text{jumlah}) \times 100\%$  ;  $f_i$  = frekuensi kelas ke  $i$  .

#### 4. Distribusi Frekuensi Kumulatif

Distribusi Frekuensi Komulatif, yaitu suatu distribusi frekuensi yang menunjukkan jumlah frekuensi berdasarkan jumlah dari masing-masing frekuensi tiap kelasnya terhadap nilai tepi kelasnya. Distribusi frekuensi komulatif dibagi menjadi dua, yaitu distribusi frekuensi komulatif kurang dari dan distribusi frekuensi komulatif lebih dari

Bentuk umum dari tabel distribusi frekuensi komulatif "kurang dari":

Nilai Data	Frekuensi Kumulatif
kurang dari a	0
kurang dari c	$f_1$
kurang dari e	$f_1 + f_2$
kurang dari g	$f_1 + f_2 + f_3$
kurang dari i	$f_1 + f_2 + f_3 + f_4$
kurang dari k	$f_1 + f_2 + f_3 + f_4 + f_5$

Bentuk umum dari tabel distribusi frekuensi komulatif "atau lebih":

Nilai Data	Frekuensi Kumulatif
a atau lebih	$f_1 + f_2 + f_3 + f_4 + f_5$
a atau lebih	$f_2 + f_3 + f_4 + f_5$
a atau lebih	$f_3 + f_4 + f_5$
a atau lebih	$f_4 + f_5$
a atau lebih	$f_5$
a atau lebih	0

#### Distribusi Frekuensi Relatif Kumulatif

Distribusi frekuensi komulatif relatif, yaitu suatu distribusi frekuensi komulatif yang frekuensi masing-masing kelasnya dinyatakan dalam bentuk prosentase.

Bentuk umum dari tabel distribusi frekuensi relatif kumulatif "kurang dari":

Nilai Data	Frekuensi Relatif Kumulatif (%)
kurang dari a	0
kurang dari c	$g_1$
kurang dari e	$g_1 + g_2$
kurang dari g	$g_1 + g_2 + g_3$
kurang dari i	$g_1 + g_2 + g_3 + g_4$
kurang dari k	100

Bentuk umum dari tabel distribusi frekuensi relatif kumulatif "atau lebih":

Nilai Data	Frekuensi Relatif Kumulatif (%)
a atau lebih	100
a atau lebih	$g_2 + g_3 + g_4 + g_5$
a atau lebih	$g_3 + g_4 + g_5$
a atau lebih	$g_4 + g_5$
a atau lebih	$g_5$
a atau lebih	0

Contoh (Data penghasilan buruh):

Kelas	Frekuensi	Frekuensi Relatif	Frekuensi Relatif Kumulatif
[40, 50)	2	0,050	0,050
[50, 60)	8	0,200	0,250
[60, 70)	17	0,425	0,625
[70, 80)	9	0,225	0,900
[80, 90)	3	0,075	0,975
[90, 100)	1	0,025	1,0

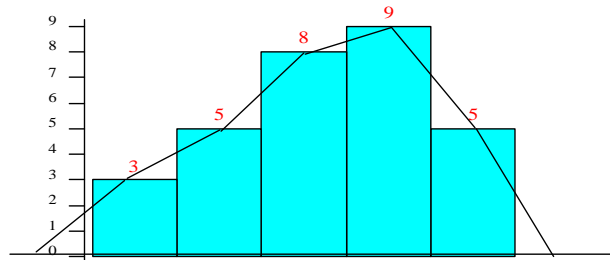
#### 4. HISTOGRAM DAN POLIGON FREKUENSI

Adalah dua gambaran secara grafik dari distribusi frekuensi.

- Histogram terdiri dari himpunan siku empat yang mempunyai :
  - Alas pada sumbu mendatar dengan pusat pada nilai tengah dan panjang sama dengan ukuran selang kelas (panjang kelas)
  - Luas sebanding terhadap frekuensi kelas.
- Poligon frekuensi adalah grafik dari frekuensi kelas yang dapat diperoleh dengan cara menghubungkan titik tengah dari puncak siku empat dalam



histogram.



Gambar 2: Histogram dan Poligon Frekuensi

### Latihan

- Data Kuantitatif

Kepala Sekolah SMA Maju berkeinginan melihat gambaran yang lebih jelas tentang distribusi penghasilan orang tua siswa. Untuk itu diambil 50 orang tua siswa sebagai sampel, kemudian dicatat penghasilan per bulannya (dalam puluhan ribu rupiah). Berikut hasilnya:

91, 78, 93, 57, 75, 52, 99, 80, 97, 62, 71, 69, 72, 89, 66, 75, 79, 75, 72, 76  
104, 74, 62, 68, 97, 105, 77, 65, 80, 109, 85, 97, 88, 68, 83, 68, 71, 69, 67, 74  
62, 82, 98, 101, 79, 105, 79, 69, 62, 73

Buatlah : Distribusi frekuensinya, distribusi frekuensi kumulatif dan distribusi frekuensi relatif serta histogram.

# MATERI KULIAH: STATISTIKA

**Bahan Kajian** : Perhitungan Ukuran Terdensi Sentral  
**Program Studi** : Pendidikan Teknik Informatika  
**Minggu ke-** : 4

## A. Pembelajaran :

### 1. Learning Outcomes

Mampu melakukan perhitungan ukuran tendensi sentral

### 2. Materi

- a) Rata-rata (mean)
- b) Median
- c) Modus

### 3. Uraian Materi

#### A. Mean

Rata-rata (mean) adalah nilai yang mewakili himpunan atau sekelompok data (*a set of data*). Nilai rata-rata umumnya cenderung terletak ditengah satu kelompok data yang disusun menurut besar kecilnya nilai. Beberapa jenis rata-rata yang sering dipergunakan adalah rata-rata hitung (mean), rata-rata ukur (geometric mean) dan rata-rata harmonic (harmonic mean).

Rata-rata hitung sebagai salah satu ukuran pemusatan mempunyai sifat-sifat sebagai berikut :

- o Setiap kelompok baik dalam bentuk skala interval maupun rasio mempunyai nilai hitung.
- o Semula nilai data harus dimasukkan ke dalam perhitungan rata-rata hitung
- o Satu kelompok baik kelas maupun satu kesatuan dalam populasi dan sampel hanya mempunyai satu rata-rata hitung
- o Rata-rata hitung untuk membandingkan karakteristik dua atau lebih populasi atau sampel

Rata-rata (mean) sering digunakan sebagai dasar perbandingan antara 2 kelompok nilai atau lebih. Misalnya ada 2 pembaca yaitu Arham dan Wildan dari Fakultas Teknik yang menempuh 5 macam mata kuliah yaitu statistik,

matematika, fisika, bahasa inggris, dan kimia. Untuk menentukan mana yang lebih pandai antara Arham dan Wildan dapat dipergunakan nilai rata-rata. Misalkan hasil ujian Arham dan Wildan adalah seperti disajikan dalam tabel di bawah ini :

No	Mata Kuliah	Hasil Ujian	
		Arham ( X )	Wildan ( Y )
1	Statistik	87	77
2	Matematika	79	65
3	Fisika	68	56
4	Bahasa Inggris	88	67
5	Kimia	78	65
	Jumlah	400	330
	Rata – rata	$400 = \frac{80}{5}$	$330 = \frac{66}{5}$

Dari nilai rata – rata tersebut dapat disimpulkan bahwa Arham lebih pandai dari pada Wildan.

### ➤ Rata-rata hitung

Apabila mempunyai nilai variabel X, sebagai hasil pengamatan atau observasi sebanyak N kali, yaitu X1, X2, . . . ., Xi, . . . ., XN, maka,

a) Rata-rata sebenarnya (populasi)

$$\begin{aligned} \mu &= \frac{1}{N} \sum_{i=1}^N X \\ &= \frac{1}{N} ( X1 + X2 + \dots + Xi + \dots + XN ) \end{aligned}$$

b) Rata-rata perkiraan (sampel)

Apabila rata-rata tersebut dihitung berdasarkan sampel sebanyak n di mana  $n \leq N$  observasi, maka rata-rata yang diperoleh disebut rata-rata perkiraan, atau rata-rata sampel, yang diberi symbol  $\bar{x}$  yang rumusnya adalah:

$$\begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n X \\ &= \frac{1}{n} (X1 + X2 + \dots + Xi + \dots + X11) \end{aligned}$$

$\bar{X}$  dibaca “X bar” yaitu symbol rata-rata.  $\hat{\mu}$  merupakan perkiraan  $\mu$ .

Contoh, berikut disajikan data penjualan suatu perusahaan selama 10 tahun.

X = hasil penjualan selama 10 tahun dalam jutaan rupiah

X1 = 50 (hasil penjualan tahu pertama)

X2 = 60 (hasil penjualan tahun kedua)

X3 = 40 (hasil penjualan tahun ketiga)

X4 = 70 (hasil penjualan tahun ke empat)

X5 = 80 (hasil penjualan tahun ke lima)

X6 = 90 (hasil penjualan tahun ke enam )

X7 = 100 (hasil penjualan tahun ke tujuh)

X8 = 65 (hasil penjualan tahun ke delapan)

X9 = 75 (hasil penjualan tahun ke sembilan)

X10 = 85 (hasil penjualan tahun ke sepuluh)

a) Hitunglah rata-rata hasil penjualan sebenarnya

b) Ambil sampel sebanyak  $n = 5$ , misalnya setelah diambil sampelnya diperoleh: X2, X4, X5, X8, X10. Hitung rata-rata perkiraan hasil penjualan pertahun.

Penyelesaian:

a) Rata-ratanya.

$$\begin{aligned}\mu &= \frac{1}{1} \sum_{i=1}^1 X1 \\ &= \frac{1}{1} (715) \\ &= 71,5\end{aligned}$$

Jadi, rata-rata hasil penjualan pertahun = Rp 71,5

b) Rata-rata perkiraan

$$\begin{aligned}\bar{X} &= \frac{1}{n} \sum_{i=1}^n X1 \\ &= \frac{1}{5} (60 + 70 + 80 + 65 + 85) \\ &= 72\end{aligned}$$

Jadi, rata-rata perkiraan hasil penjualan per tahun = Rp 72 juta (ternyata sangat mendekati rata-rata sebenarnya).  $\bar{x}$  merupakan perkiraan  $\mu$ .

➤ **Rata-rata hitung data berkelompok**

Apabila data sudah disajikan dalam bentuk tabel frekuensi, dimana  $X_1$  terjadi  $f_1$  kali,  $X_2$  terjadi  $f_2$  kali, dan seterusnya sampai  $X_k$  terjadi  $f_k$  kali, maka rumus rata-rata dari data yang telah dibuat table frekuensi adalah seagai berikut:

$$\bar{x} = \frac{\sum_{i=1}^k f X_i}{\sum_{i=1}^k f}$$

Karena  $\sum_{i=1}^k f_i = n$ , maka :  $\bar{x} = \frac{1}{n} \sum_{i=1}^k f_i X_i$

Atau 
$$\bar{x} = \frac{\sum_{i=1}^k M_i f_i}{\sum_{i=1}^k f_i}$$

Dimana  $M_i$  = nilai tengah kelas interval ke= $i$ . (untuk data berkelompok).

Contoh: perhatikan tabel berikut, berdasarkan data tersebut hitunglah rata-ratanya:

X	8	6	4	5	7	9	X = 39
f	2	3	4	3	2	1	f = 15
X f	16	18	16	15	14	9	X f = 88

Penyelesaian

$$\bar{x} = \frac{\sum f X}{\sum f} = \frac{88}{15} = 5,867$$

Jadi, rata-raa dari data di atas adalah 5,867

Cara lain untuk menghitung rata-rata data berkelompok yaitu dengan menggunakan rumus:

$$\bar{x} = x_0 + p \left( \frac{\sum_{i=1}^k f_i c_i}{\sum_{i=1}^k f_i} \right)$$

Contoh : Dibawah ini ada data sepuluh produk makanan dalam ratusan ribu rupiah

No	1	2	3	4	5	6	7	8	9	10	total
data	12	10	13	15	19	14	14	17	16	20	150

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{150}{10} = 15$$

Maka rata-rata penjualan 10 produk makanan = Rp. 1.500.000

Rata-rata dengan data berkelompok

Dibawah ini adalah tabel distribusi frekuensi modal dalam jutaan rupiah dari 10 perusahaan makanan.

Kelas	$f_i$	$x_i$	$f_i x_i$
60 - 65	1	62,5	62,5
66 - 71	4	68,5	274
72 - 77	1	74,5	74,5
78 - 83	3	80,5	242
84 - 89	1	86,5	86,5
Jumlah	10		739

$$\sum_{i=1}^k f_i = 10$$

$$\sum_{i=1}^k f_i x_i = 739$$

$$\bar{x} = \frac{\sum_{i=1}^k f_i x_i}{\sum_{i=1}^k f_i} = \frac{739}{10} = 73,9$$

Maka rata-rata modal 10 perusahaan makanan tersebut = Rp. 73.900.000,-

Atau dengan :

Kelas	$f_i$	$x_i$	$c_i$	$f_i c_i$
60 - 65	1	62,5	-2	-2
66 - 71	4	68,5	-1	-4
72 - 77	1	74,5	0	0
78 - 83	3	80,5	1	3
84 - 89	1	86,5	2	2
Jumlah	10			-1

$$\sum_{i=1}^k f_i = 10$$

$$x_0 = 74,5 \Rightarrow \text{Kelas interval 3}$$



$$\sum_{i=1}^k f_i c_i = -1$$

$$\bar{x} = x_0 + p \left( \frac{\sum_{i=1}^k f_i c_i}{\sum_{i=1}^k f_i} \right) = 74,5 + 6 \left( \frac{-1}{10} \right) = 74,5 + 6(-0,1)$$

$$\bar{x} = 74,5 - 0,6 = 73,9$$

$p = \text{panjang kelas} = 6$

### ➤ Rata-rata hitung Tertimbang

Sering kali dalam suatu persoalan, masing-masing nilai mempunyai nilai bobot / timbangan tertentu, misalnya  $X_1$  dengan timbangan  $W_1$ ,  $X_2$  dengan timbangan  $W_2$ , dan seterusnya sampai  $X_n$  dengan timbangan  $W_n$ . Oleh karena itu, rata-rata yang menggunakan timbangan tersebut disebut *rata-rata timbang* (*weight arithmetic mean*) dengan rumus sebagai berikut:

$$\bar{x} = \frac{\sum W X}{\sum W} = \frac{W_1 X_1 + W_2 X_2 + \dots + W_i X_i + \dots + W_k X_k}{W_1 + W_2 + \dots + W_1 + \dots + W_k}$$

Perhatikan bahwa dalam rumus diatas, timbangannya berupa frekuensi ( $W_i = f_i$ ).

Contoh: Data berikut menunjukkan nilai hasil ujian statistika mahasiswa semester III 2014 jurusan Kesejahteraan Keluarga

70, 75, 45, 70, 50, 60, 90, 75, 45, 60,

60, 90, 70, 50, 45, 75, 70, 60, 90, 45,

90, 70, 75, 60, 60, 45, 50, 60.

Carilah rata-rata nilai matematika mahasiswa dengan cara:

- a) Data tidak dikelompokkan
- b) Data dikelompokkan.

Penyelesaian :

$$\begin{aligned}
 \text{a) } \bar{x} &= \frac{\sum X}{n} \\
 &= \frac{1}{2} (70 + 75 + \dots + 60) \\
 &= 64,46
 \end{aligned}$$

b)

X = (nilai statistika)	45	50	60	70	75	90
f = (banyak mahasiswa)	5	3	7	5	4	4

$$\begin{aligned}
 \bar{x} &= \frac{\sum f X}{\sum f} \\
 &= \frac{45 (5) + 50 (3) + 60 (7) + 70 (5) + 75(4) + 90 (4)}{5+3+7+5+4+4} \\
 &= 64,46
 \end{aligned}$$

## B. Median

Median adalah nilai tengah dari kumpulan data yang telah diurutkan (disusun) dari data terkecil sampai data terbesar. Median dari suatu himpunan bilangan yang disusun menurut urutan besarnya (yaitu dalam suatu array) adalah nilai pertengahan atau nilai tengah hitung dari perhitungan.

Median adalah nilai tengah dari data yang ada setelah data yang diurutkan. Median merupakan rata – rata apabila ditinjau dari segi kedudukannya dalam urutan data. Median sering pula disebut rata – rata posisi. Median ditulis singkat atau disimbolkan dengan Me atau Md. Cara mencari median dibedakan antara data tunggal dan data berkelompok.

### 1). Median data tunggal

Median untuk data tunggal dapat dicari dengan pedoman sebagai berikut.

) Jika jumlah data ganjil, mediannya adalah data yang berada paling tengah.



J) Jika jumlah data genap, mediannya adalah hasil bagi jumlah dua data yang berada di tengah. Pedoman tersebut dirumuskan sebagai berikut.

a) Untuk data ganjil( n = ganjil )

$$Me = X_{\frac{n+1}{2}}$$

b) Untuk data genap (n = genap)

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

Atau secara singkat median dapat ditentukan :  $Me = \text{nilai yang ke} \frac{1}{2}(n + 1)$

ContohSoal :

Tentukan Median dari data berikut !

a. 4, 3, 2, 6, 7, 5, 8

b. 11, 5, 7, 4, 8, 14, 9, 12

**Penyelesaian :**

a. Urutan data : 2, 3, 4, 5, 6, 7, 8

Jumlah data ( n ) = 7 ( ganjil )

$$Me = \frac{X_{7+1}}{2} = X_4 = 5$$

b. Urutan data : 4, 5, 7, 8, 9, 11, 12, 14

Jumlah data ( n ) = 8 ( genap )

$$Me = \frac{X_4}{2} + \frac{X_5}{2} = \frac{8}{2} + \frac{9}{2} = 8,5$$

## 2). Median data berkelompok

Median untuk data berkelompok dapat dicari dengan rumus sebagai berikut.

$$Me = \frac{B + \frac{1}{2}n - (\sum f_2)^o \cdot C}{f_{Me}}$$

Keterangan :

Me = Median

B = tepi bawah kelas median

$n$  = jumlah frekuensi

$(\sum f_2)^0$  = jumlah frekuensi kelas-kelas sebelum kelas median

$C$  = panjang interval kelas

$f_{Me}$  = frekuensi kelas median

Dalam mencari media data kelompok (distribusi frekuensi) yang perlu dicari terlebih dahulu adalah kelas tempat median berada (kelas median). Kelas media dapat dicari dengan :  $(\sum f_2)^0 \frac{1}{2}n$ .

Contoh Soal :

Tentukan median dari distribusi frekuensi berikut !

Diameter dari 40 cake

Diameter cake ( mm )	Frekuensi ( $f$ )
65 – 67	2
68 – 70	5
71 – 73	13
74 – 76	14
77 – 79	4
80 – 82	2
Jumlah	40

Penyelesaian :

Jumlah frekuensi (  $n$  ) = 40 dan  $\frac{1}{2}n = 20$

Kelas median adalah  $(\sum f_2)^0 \frac{1}{2}n$

$$f_1 + f_2 + f_3 = 20 \quad 20$$

Jadi, kelas median adalah kelas ke – 3

$$B = 70,5$$

$$(\sum f_2)^0 = 7$$

$$C = 3$$

$$f_{Me} = 13$$

$$\begin{aligned}
 Me &= B + \frac{\frac{1}{2}n - (\sum f_2)0}{f_{Me}} \cdot C \\
 &= 70,5 + \frac{\frac{2}{1} - 7}{1} \times 3 \\
 &= 73,5
 \end{aligned}$$

### C. Modus

Modus adalah nilai data yang paling sering muncul atau nilai data yang frekuensinya paling besar. Data yang belum dikelompokkan bisa memiliki satu modus, dua modus, atau mungkin tidak mempunyai modus. Data yang memiliki satu modus disebut mono modus, sedangkan data yang memiliki dua modus disebut bimodus.

#### 1. Modus dari data tunggal

Contoh :

Tentukan modus dari data dibawah ini !

5,7,7,6,8,6,6,5,8,6

Jawab :

Setelah data diurutkan diperoleh : 5,5,6,6,6,6,7,7,8,8

Modus ( $M_o$ ) = 6

#### 2. Modus dari data yang dikelompokkan

Untuk menghitung modus dari data yang telah dikelompokkan dipergunakan rumus sebagai berikut :

$$Mo = b + P \times \frac{b_1}{b_1 + b_2}$$

Keterangan :

$M_o$  = modus

$b$  = batas bawah kelas

$p$  = panjang kelas

$b_1$  = frekuensi kelas modus dikurangi frekuensi kelas sebelumnya

$b_2$  = frekuensi kelas modus dikurangi frekuensi kelas berikutnya

Contoh : Tentukan Modus dari data sebagai berikut !

Nilai Frekuensi

Nilai	Frekuensi
52-58	2
59-65	6
66-72	7
73-79	20
80-86	8
87-93	4
94-100	3
Jumlah	50

Jawab :

Frekuensi terbanyak pada kelas 73 – 79, berarti modusnya terletak pada kelas 73 – 79.

$$b = \frac{72 + 73}{2} = 72,5$$

$$P = 7 ; b_1 = 20 - 7 = 13 \text{ dan } b_2 = 20 - 8 = 12$$

$$M_o = b + P \frac{b_1}{b_1 + b_2}$$

$$M_o = 72,5 + 7 \frac{13}{13 + 12} = 72,5 + 7 \frac{13}{25} = 72,5 + 3,64$$

$$M_o = 76,14$$

Jadi, modus adalah 76,14

# MATERI KULIAH: STATISTIKA

**Bahan Kajian** : Perhitungan Ukuran Sebaran  
**Program Studi** : Pendidikan Teknik Informatika  
**Minggu ke-** : 5

**A. Pembelajaran :**

**1. Learning Outcomes**

Memiliki kemampuan menentukan ukuran sebaran

**2. Materi**

- a) Rentang
- b) Deviasi rata-rata
- c) Varian
- d) Deviasi standar

**3. Uraian Materi**

**a. Rentang (Rentang)**

Range = selisih nilai data terbesar – data terkecil = L – S

L = nilai data terbesar

S = nilai data terkecil

Contoh data:

44 56 60 67 70 80 85 90 99

Range = 99 – 44 = 55.

**b. Deviasi Rata-Rata (Rata-Rata Simpangan)**

Jumlah semua jarak antara tiap data dengan rata-rata dibagi banyaknya data

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

Contoh: 4, 5, 7, 8, 8, 10 ( n = 6 dan  $\bar{x} = 7$  )

$$\text{maka RS} = \frac{-4 - 7 + -5 - 7 + \dots -10 - 7}{6} = 1.67$$

Contoh :

$x_i$	$x_i - \bar{x}$	$ x_i - \bar{x} $
8	- 1	1
7	- 2	2
10	1	1
11	2	2

Dari data di samping ini, jika dihitung, rata-ratanya =  
 9. Jumlah harga-harga mutlakanya, yaitu jumlah  
 bilangan-bilangan pada kolom akhir, adalah 6. Maka

$$RS = \frac{6}{4} = 1 \frac{1}{2}$$

Untuk data berkelompok

$$RS = X \frac{\sum f |m - \bar{x}|}{n}$$

- dengan :
- $m$  = titik tengah kelas
  - $f$  = frekuensi tiap kelas interval
  - $\bar{x}$  = rata-rata hitung
  - $n$  = banyak data/jumlah frekuensi

Contoh:

Hitung rata-rata simpangan dari data dibawah ini:

Kelas ke	Interval	f
1	160 - 303	2
2	304 - 447	5
3	448 - 591	9
4	592 - 735	3
5	736 - 878	1

Jawab

$$\bar{x} = \frac{\sum_{i=1}^n fx_i}{n} = \frac{9813,5}{20} = 490,7$$

Kelas ke	Interval	m	f	fx	$ mZ\bar{x} $	$f    mZ\bar{x} $
1	160 - 303	231,5	2	463	259,2	518,4
2	304 - 447	375,5	5	1877,5	115,2	576
3	448 - 591	519,5	9	4675,5	28,8	259,2
4	592 - 735	663,5	3	1990,5	172,8	518,4
5	736 - 878	807	1	807	316,3	316,3
Jumlah		2597	20	9813,5	892,3	2188,3

$$RS X = \frac{f | |mZ\bar{x}|}{n}$$

$$RS = 2188,3/20 = 109,415$$

Contoh

Untuk data pengeluaran per hari 30 keluarga (yang telah dikelompokan),

hitunglah simpangan rata-ratanya !

Pengeluaran (ribu Rp)	f	m	$ mZ\bar{x} $	$f    mZ\bar{x} $
50 – 55	1	52,5		
56 – 61	5	58,5		
62 – 67	6	64,5		
68 – 73	10	70,5		
74 – 79	5	76,5		
80 – 85	3	82,5		
Jumlah	30	-		

**c. Varian**

Rumus yang dipergunakan untuk menghitung varians, jika data berasal dari populasi adalah:

$$s^2 = \frac{\sum_{i=1}^N X_i^2 - \frac{(\sum_{i=1}^N X_i)^2}{N}}{N}$$

Sedangkan varians yang dihitung berdasarkan sampel dihitung dengan rumus:

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

Varians dan simpangan baku hanya boleh digunakan sebagai alat pembanding keseragaman data, apabila data yang dibandingkan keseragamannya itu berasal dari variabel yang sama dengan satuan pengukuran yang sama pula. Varians dan simpangan baku hanya valid digunakan sebagai ukuran variasi untuk variabel yang memenuhi tingkat pengukuran sekurang-kurangnya interval.

- o Bentuk lain untuk rumus *varians* ialah :  $s^2 = \frac{n \sum x_i^2 - (\sum x_i)^2}{n(n-1)}$

Pada rumus ini tidak perlu dihitung rata-rata.

$x_i$	$x_i^2$
8	64
7	49
10	100
11	121
4	16
40	350

Dihasilkan  $\sum x_i = 40$  dan  $\sum x_i^2 = 350$ .

Dengan  $n = 5$ , didapat varians

$$s^2 = \frac{5 \times 350 - (40)^2}{5 \times 4} = 7,5 \text{ dan}$$

$$s = \sqrt{7,5} = 2,74.$$



- o Untuk data dari sampel telah disusun dalam daftar distribusi frekuensi, varians  $s^2$  dipakai rumus :

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

Untuk:

$x_i$  = tanda kelas,

$f_i$  = frekuensi yang sesuai dengan tanda kelas  $x_i$

$n = \sum f_i$ .

Simpangan Baku dan Varian untuk data berkelompok adalah :

$$s^2 = \frac{\sum f m^2 Z^2 - \frac{(\sum f m Z)^2}{n}}{n}$$

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

Contoh

Hitunglah Varians dan Standar Deviasi untuk pengeluaran per hari 30 keluarga yang telah dikelompokan !

Pengeluaran (ribu Rp)	$f$	$m$	$mZ\bar{x}$	$f m^2 Z^2$	$f   f m Z \bar{x}$
50 – 55	1	52,5			
56 – 61	5	58,5			
62 – 67	6	64,5			
68 – 73	10	70,5			
74 – 79	5	76,5			
80 – 85	3	82,5			
Jumlah	<b>30</b>	-			

#### d. Devisi Standar

Yang dimaksud dengan “simpangan baku atau deviasi standar” adalah suatu nilai yang menunjukkan besarnya simpangan rata-rata seluruh nilai yang ada

dalam kelompok data dengan nilai pusatnya dengan cara menghilangkan kemungkinan nilai nol dengan jalan dikuadratkan. Deviasi standar dari suatu rangkaian data adalah akar pangkat dua dari rata-rata kudarat selisih nilai data individual terhadap mean rangkaian data itu. Terdapat dua jenis rumus yang umum digunakan untuk deviasi standar, yaitu: deviasi standar untuk populasi dan diberi simbol  $\Sigma$  (baca : sigma), deviasi standar untuk sampel disimbul dengan s. Oleh karena itu, kita harus memilih rumus yang sesuai dengan jenis data yang ada, yaitu data populasi atau data sampel.

Jika data kita adalah data populasi gunakan rumus deviasi standar untuk populasi, dan jika data kita adalah data sampel, maka gunakan rumus deviasi standar untuk sampel.

Jika kita mempunyai sampel berukuran n dengan data  $x_1, x_2, \dots, x_n$  dan

rata-rata  $\bar{x}$ , maka statistik s dihitung dengan:  $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$  atau

<p>Rumus deviasi standar untuk populasi</p> $\Sigma \sqrt{\frac{\sum (x - \bar{x})^2}{N}}$ <p>Keterangan:</p> <p><math>\Sigma</math> = standar deviasi populasi</p> <p>x = nilai pengamatan</p> <p><math>\bar{x}</math> = mean populasi</p> <p>N = jumlah pengamatan dalam populasi</p>	<p>Rumus deviasi standar untuk sampel</p> $s \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$ <p>Keterangan:</p> <p>s = standar deviasi sampel</p> <p>x = nilai pengamatan</p> <p><math>\bar{x}</math> = mean sampel</p> <p>n = jumlah pengamatan dalam sampel</p>
---	--

Pangkat dua dari simpangan baku dinamakan *varians*.

o Simpangan baku s dihitung sebagai berikut

- 1). Hitung rata-rata  $\bar{x}$
- 2). Tentukan selisih  $x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}$

- 3). Tentukan kuadrat selisih tersebut, yakni  $(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \dots, (x_n - \bar{x})^2$
- 4). Kuadrat-kuadrat tersebut dijumlahkan
- 5). Jumlah tersebut dibagi oleh  $(n - 1)$
- 6). Lalu diambil akarnya yang positif.

Contoh

Diberikan sampel dengan data : 8, 7, 10, 11, 4.

Untuk menentukan simpangan baku  $s$ , kita buat tabel berikut:

$x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
8	0	0
7	- 1	1
10	2	4
11	3	9
4	- 4	16

Rata-rata  $\bar{x} = 8$ , dari kolom (2), bahwa

$\sum (x_i - \bar{x}) = 0$ . Karena itulah di sini diambil kuadratnya yang dituliskan pada kolom (3). Didapat

$$\sum (x_i - \bar{x})^2 = 30.$$

$$\text{didapat : } S = \sqrt{\frac{30}{40}} \times \sqrt{7,5} = 2,74.$$

# MATERI KULIAH: STATISTIKA

**Bahan Kajian** : Hipotesis  
**Program Studi** : Pendidikan Teknik Informatika  
**Minggu ke-** : 6

## A. Pembelajaran :

### 1. Learning Outcomes

Memiliki pemahaman yang luas, cerdas dan teliti tentang hipotesis. Pengenalan Program SPSS

### 2. Materi

- a) Pengertian hipotesis
- b) Jenis hipotesis
- c) SPSS

### 3. Uraian Materi

#### A. Pengertian Hipotesis

Dalam statistika, pengujian hipotesis merupakan bagian terpenting untuk mengambil keputusan. Dengan melakukan pengujian hipotesis seorang peneliti akan mendapat jawaban pertanyaan-pertanyaan yang diajukannya dengan menyatakan penolakan atau penerimaan terhadap hipotesis.

Hipotesis adalah jawaban sementara terhadap masalah penelitian yang harus diuji atau pernyataan yang diterima secara sementara yang didasarkan pada hasil studi literatur. Hipotesis biasanya memuat pernyataan-pernyataan yang bersifat netral atau hal yang umum terjadi. Kebenaran hipotesis secara pasti tidak pernah diketahui kecuali jika dilakukan pengamatan terhadap seluruh anggota populasi. Untuk melakukan hal ini tidaklah efisien apabila ukuran populasinya sangat besar. Secara statistik Hipotesis menyatakan parameter populasi dari suatu variabel yang terdapat dalam populasi dan dihitung berdasarkan statistik sampel. Karena merupakan dugaan sementara, maka hipotesis mungkin benar, tetapi mungkin juga tidak benar

Suatu prosedur pengujian hipotesis tentang parameter populasi menggunakan informasi dari sampel dan teori probabilitas untuk menentukan apakah hipotesis tersebut secara statistik dapat diterima atau ditolak. Tujuan pengujian hipotesis adalah ingin mendapatkan kesimpulan mengenai suatu

populasi berdasarkan sampel yang kita miliki. Bila ingin mengetahui pendapat mahasiswa tentang khasiat tempuyung sebagai diuretik dan menanyakan kepada seluruh mahasiswa → observasi → analisis deskriptif → tidak perlu uji hipotesis. Tetapi bila kita hanya mengambil sampel mahasiswa → uji hipotesis → untuk membuktikan jawaban dari sampel bisa mewakili jawaban seluruh mahasiswa. Kesimpulan dari pengujian hipotesis secara statistik hanya berupa menerima atau menolak hipotesis dan ini tidak membuktikan kebenaran hipotesis karena statistika sama sekali tidak melakukan pembuktian. Penerimaan suatu hipotesis terjadi karena tidak cukup bukti untuk menolak hipotesis tersebut dan bukan karena hipotesis itu benar.

Penarikan sejumlah contoh acak dari suatu populasi, diamati karakteristiknya dan kemudian dibandingkan dengan hipotesis yang diajukan merupakan suatu langkah melakukan uji hipotesis. Apabila contoh acak ini memberikan indikasi atau petunjuk yang mendukung hipotesis yang diajukan maka hipotesis tersebut diterima, sedangkan apabila contoh acak itu memberikan indikasi yang bertentangan dengan hipotesis yang diajukan, maka hipotesis tersebut ditolak.

#### Kegunaan pengujian hipotesis:

1. Memberikan batasan/memperkecil jangkauan kerja penelitian
2. Memperhatikan kondisi fakta dan hubungan antar fakta
3. Memfokuskan fakta
4. Panduan pengujian

#### Ciri Hipotesis

1. Menyatakan hubungan
  2. Sesuai dengan fakta
  3. Berhubungan dengan ilmu
  4. Dapat diuji
  5. Sederhana
- Bisa menerangkan fakta

## B. Jenis Hipotesis

### Jenis Hipotesis:

1. Hipotesis deskriptif : Rata-rata situasi kepemimpinan paling rendah 40% dari yang diharapkan
2. Hipotesis assosiatif : Terdapat hubungan yang positif dan signifikan antara gaya kepemimpinan dengan iklim kerja organisasi
3. Hipotesis komparatif : Terdapat perbedaan gaya kepemimpinan yang signifikan antara pimpinan eselon II, III, IV

Prosedur untuk menguji suatu hipotesis adalah sebagai berikut :

1. Rumuskan hipotesis nol dan hipotesis alternative.

Ada 3 bentuk rumusan hipotesis yaitu deskriptif, hubungan dan komparatif.

*Hipotesis nol* ( $H_0$ ) adalah suatu pernyataan tentang populasi. Pada hipotesis nol biasanya ada istilah *tidak*, seperti tidak ada perbedaan secara nyata ataupun tidak ada perubahan secara nyata. Sebagai contoh untuk pernyataan rata-rata keuntungan perusahaan per bulan dari suatu perusahaan makanan adalah Rp 200.625.000, maka hipotesis nol ditulis sebagai  $H_0 : \hat{\mu} = 200.625.000$ . Contoh lain pernyataan tidak ada perbedaan nyata antara proporsi konsumen yang membeli rendang dengan proporsi konsumen yang membeli kalio daging, maka hipotesis nol ditulis sebagai  $H_0 : p_{rendang} = p_{kalio\ daging}$ .

*Hipotesis alternatif* menyatakan apa yang akan disimpulkan sekiranya hipotesis nol ditolak. Sehingga nilai sama dengan (=) tidak akan pernah ada dalam hipotesis alternatif. Hipotesis alternatif hanya akan dibuktikan jika hipotesis nol adalah tidak benar. Sebagai contoh hipotesis alternatif untuk pernyataan rata-rata keuntungan perusahaan per bulan dari suatu perusahaan makanan adalah tidak sama dengan Rp 200.625.000, maka hipotesis alternatif ditulis sebagai  $H_1 : \hat{\mu} \neq 200.625.000$ .

Jika pada hipotesis alternatif tidak ada arah penunjuk (apakah lebih besar atau kecil, hanya ada tanda tidak sama dengan) maka dilakukan *uji dua arah*. Jika

hipotesis alternatif menyatakan suatu arah ( $<$  atau  $>$ ), maka uji yang dilakukan adalah uji *satu arah*.

2. Pilih suatu taraf nyata/signifikansi

Taraf signifikansi ini sering juga disebut tingkat resiko, yaitu resiko yang diambil untuk menolak hipotesis nol apabila hipotesis nol adalah benar. Seorang peneliti perlu menetapkan taraf signifikansi sebelum menentukan aturan pengambilan keputusan dan mengumpulkan data sampel. Taraf signifikansi yang diambil tergantung pada jenis penelitian yang dilakukan. Sebagai contoh, biasanya dalam penelitian pangan, taraf signifikansi adalah 0,05 atau untuk pengendalian mutu biasanya diambil taraf signifikansi 0,01. Taraf signifikansi biasa diberi tanda  $\alpha$ . Taraf signifikansi ini juga sering disebut sebagai peluang melakukan Kesalahan Jenis I dalam pengujian hipotesis. Terdapat jenis kesalahan lain dalam pengujian hipotesis yaitu jika jenis kesalahan II, dimana hipotesis nol diterima padahal hipotesis nol tersebut salah. Peluang melakukan Kesalahan Jenis II dalam pengujian hipotesis disebut  $\beta$ .

Secara lebih jelas, keputusan yang dapat diambil oleh peneliti dan kemungkinan konsekuensinya ditunjukkan pada tabel 7.

Tabel 7. Keputusan dan konsekuensi dalam pengujian hipotesis

Hipotesis nol	Peneliti	
	Menerima $H_0$	Menolak $H_0$
$H_0$ benar	Keputusan tepat	Kekeliruan Tipe I ( $\alpha$ )
$H_0$ salah	Kekeliruan Tipe II ( $\beta$ )	Keputusan tepat

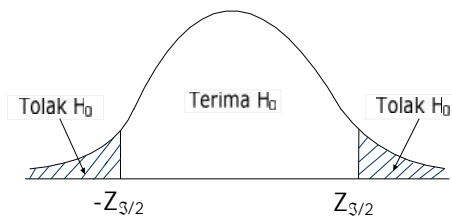
3. Tentukan uji statistik

Uji statistik adalah suatu nilai yang diperoleh dari nilai statistik sampel, yang akan digunakan untuk menerima atau menolak hipotesis. Terdapat berbagai uji statistik antara lain dikenal Z, F, t dan  $\chi^2$  (kai kuadrat sebagai uji statistik).

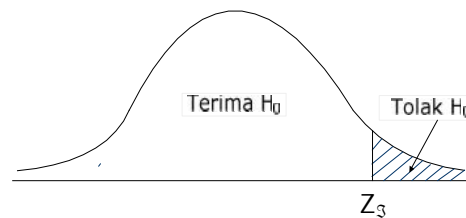
4. Buat aturan pengambilan keputusan

Aturan keputusan merupakan pernyataan mengenai kondisi dimana hipotesis nol ditolak dan kondisi dimana hipotesis nol diterima. Aturan keputusan ini biasa digambarkan dalam suatu gambar distribusi yang bersesuaian. Langkah ini bergantung pada langkah-langkah sebelumnya.

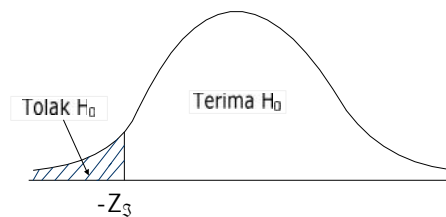
- a. Jika uji yang dilakukan adalah uji dua arah, taraf signifikansi  $\alpha$  dengan uji statistik  $Z$ , maka gambar aturan keputusannya adalah :



- b. Jika uji yang dilakukan adalah uji satu arah (dengan tanda pada hipotesis alternative  $>$ ), taraf signifikansi  $\alpha$  dengan uji statistik  $Z$ , maka gambar aturan keputusannya adalah :



- c. Jika uji yang dilakukan adalah uji satu arah (dengan tanda pada hipotesis alternative  $<$ ), taraf signifikansi  $\alpha$  dengan uji statistik  $Z$ , maka gambar aturan keputusannya adalah :



Nilai  $Z_{\alpha/2}$ ,  $Z_{\alpha}$  atau  $-Z_{\alpha}$  disebut dengan nilai kritis, yaitu titik yang memisahkan daerah dimana hipotesis nol ditolak dan diterima. Nilai ini



didapat dari tabel distribusi yang bersesuaian. Jika Z didapat dari tabel distribusi normal standar.

5. Membuat keputusan dan kesimpulan

Keputusan dibuat dengan melihat apakah nilai uji statistik berada di daerah penolakan atau penerimaan  $H_0$ . Setelah diputuskan  $H_0$  ditolak atau diterima, maka kesimpulan akan dibuat berdasarkan permasalahan yang ada. Perlu juga ditekankan taraf signifikansi pada saat membuat kesimpulan.

a) Hipotesis Deskriptif

- ❖ Merupakan hipotesis tentang nilai suatu variabel mandiri, tidak membuat perbandingan atau hubungan.
- ❖ Sebagai contoh bila rumusan masalah penelitian sbb:  
Seberapa tinggi kandungan parasetamol dalam tablet x?  
Berapa lama *expired date* produk A pada T refri?

**Rumusan hipotesis:**

- Kandungan parasetamol dalam tablet X adalah 500 mg/tablet.
- Expired date produk X A pada suhu refri 10 hari.

b) Hipotesis Asosiatif

- ❑ Pernyataan yang menunjukkan dugaan tentang hubungan antara dua variabel atau lebih.
- ❑ Sebagai contoh rumusan hipotesis asosiatif:
  - Apakah ada hubungan antara dosis obat dengan aktivitas?
  - Apakah ada pengaruh penambahan CMC terhadap viskositas larutan?

**Rumusan hipotesis:**

- Tidak ada hubungan antara jumlah iklan dengan volume penjualan.  $H_0: \rho = 0$   $H_a: \rho \neq 0$
- Tidak ada pengaruh penambahan CMC terhadap viscositas sari buah tomat .  $H_0: \rho = 0$   $H_a: \rho \neq 0$ .

c) Hipotesis komperatif

- Pernyataan yang menunjukkan dugaan nilai dalam satu variabel atau lebih pada sampel yang berbeda.
- Sebagai contoh rumusan hipotesis komparatif:  
Apakah ada perbedaan roti berbahan terigu dengan roti yang disubtitusi tapioka X? Apakah ada perbedaan viskositas larutan A dan B?

**Rumusan hipotesis:**

- ❖ Tidak ada perbedaan kandungan roti berbahan terigu (1) dengan roti yang disubtitusi tapioka (2).  $H_0: \hat{\mu}_1 = \hat{\mu}_2$   $H_a: \hat{\mu}_1 \neq \hat{\mu}_2$
- ❖ Viscositas larutan A tidak berbeda dibandingkan larutan B.  $H_0: \hat{\mu}_1 = \hat{\mu}_2$   
 $H_a: \hat{\mu}_1 \neq \hat{\mu}_2$ .

Pada bab ini akan dibahas pengujian hipotesis untuk satu sampel dan untuk dua sampel.

**A. UJI HIPOTESIS UNTUK SATU SAMPEL**

**1. Uji Z untuk Pengujian Hipotesis Rata-Rata ( $\exists$  diketahui)**

Pernyataan-pernyataan sebagai berikut ini melibatkan rata-rata populasi :

- a. Rata-rata umur dosen di Universitas Negeri Padang kurang dari 60 tahun
- b. Rata-rata penghasilan perusahaan lebih dari Rp 15 juta per bulan di Padang
- c. Rata-rata tingkat efisiensi pegawai bagian produksi sama dengan 200

Sekiranya pernyataan-pernyataan tersebut perlu diuji, maka pernyataan pertama dan kedua diuji dengan uji satu arah dan pernyataan ketiga dengan uji dua arah. Jika simpangan baku populasi diketahui, maka statistik uji yang digunakan

adalah: 
$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Pengujian hipotesis dengan uji Z ini beserta langkah-langkahnya akan diterangkan pada contoh-contoh berikut.

### Contoh

Restoran MacBurger menyatakan bahwa rata-rata hitung waktu menunggu bagi pelanggan untuk mendapat pelayanan berdistribusi normal dengan rata-rata waktu menunggu 3 menit dan simpangan baku 1 menit. Untuk mengendalikan mutu pelayanannya, 50 pelanggan dipilih dan didapat rata-rata waktu menunggu adalah 2,75 menit. Dapatkan bagian pengendalian mutu menyimpulkan bahwa rata-rata waktu menunggu kurang dari 3 menit pada taraf signifikan 0,05?

Jawab :

**Langkah 1.** Kalimat “Dapatkan bagian pengendalian mutu menyimpulkan bahwa rata-rata waktu menunggu kurang dari 3 menit?” menunjukkan bahwa Hipotesis nol adalah Rata-rata populasi lebih atau sama dengan 3 menit dan hipotesis alternatif adalah Rata-rata populasi kurang dari 3 menit. Kedua hipotesis ditulis sebagai berikut :

$$H_0 : \mu \geq 3$$

$$H_1 : \mu < 3$$

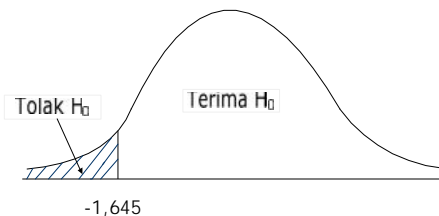
Ini merupakan uji satu arah karena hipotesis alternative menyatakan arah perbedaan yaitu *kurang dari*.

**Langkah 2.** Taraf signifikansi adalah 0,05.

**Langkah 3.** Uji statistik adalah Z karena simpangan baku populasi ( $\sigma$ ) diketahui sebesar 1, dan dari sampel dengan  $n=50$  diperoleh rata-rata waktu menunggu sampel  $\bar{x} = 2,75$  maka :

$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{2,75 - 3}{\frac{1}{\sqrt{50}}} = -1,77$$

**Langkah 4.** Uji yang dilakukan adalah uji satu arah, taraf signifikansi  $\alpha=0,05$  maka nilai kritis  $-Z_{\alpha} = -1,645$ , gambar aturan keputusannya adalah :



**Langkah 5.** Karena  $Z=-1,77$  berada di daerah Tolak  $H_0$ , maka keputusannya adalah menolak  $H_0$ . Kesimpulannya pada taraf signifikan 0,05, bagian pengendalian mutu Restoran MacBurger bisa menyimpulkan bahwa rata-rata waktu menunggu adalah kurang dari 3 menit.

**2. Uji t untuk Pengujian Hipotesis Rata-Rata ( $\exists$  tidak diketahui)**

Pada kebanyakan pengujian hipotesis, nilai simpangan baku populasi adalah tidak diketahui. Sebagai gantinya nilai simpangan baku populasi ditaksir oleh simpangan baku sampel. Jika populasi diasumsikan berdistribusi normal, maka distribusi samplingnya akan berdistribusi t dengan derajat kebebasan  $n-1$ . Sehingga statistik uji untuk pengujian hipotesis rata-rata dengan simpangan baku populasi tidak diketahui adalah

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Langkah-langkah pengujian hipotesisnya sama dengan yang telah diterangkan pada bagian sebelumnya. Pada langkah 4, nilai kritis ditentukan oleh nilai pada tabel distribusi t dengan derajat kebebasan  $n-1$ .

**Contoh .**

Dalam waktu lima tahun terakhir, rata-rata invois penjualan di PT A adalah Rp 1,2 juta. Akuntan di perusahaan tersebut perlu menginformasikan kepada bagian keuangan sekiranya terjadi perubahan pada rata-rata invois penjualan tersebut. Untuk tujuan itu, maka akuntan tersebut memilih 12 invois penjualan (dalam jutaan Rp) secara acak seperti berikut :

1,08 1,52 1,11 1,10 1,27 1,07  
0,93 0,91 1,11 0,75 1,28 1,35

Pada taraf signifikan 0,05 , apakah akuntan tersebut perlu melaporkan pada bagian keuangan bahwa telah terjadi perubahan pada rata-rata invois penjualan?

Jawab :

**Langkah 1.** Kalimat “rata-rata invoice penjualan di PT A adalah Rp 1,2 juta” dan “Apakah telah terjadi perubahan pada rata-rata invoice penjualan” menunjukkan bahwa Hipotesis nol adalah Rata-rata populasi sama dengan 1,2 dan hipotesis alternatif adalah Rata-rata populasi tidak sama dengan 1,2. Kedua hipotesis ditulis sebagai berikut :

$$H_0 : \mu = 1,2$$

$$H_1 : \mu \neq 1,2$$

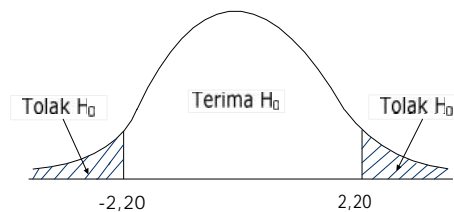
Ini merupakan uji dua arah karena hipotesis alternative tidak menyatakan arah perubahan.

**Langkah 2.** Taraf signifikansi adalah 0,05.

**Langkah 3.** Uji statistik adalah t karena simpangan baku populasi ( $\sigma$ ) tidak diketahui. Dari sampel dengan  $n=12$  dapat dihitung rata-rata invoice penjualan  $\bar{x} = 1,12$  dan  $s=0,21$  maka :

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{1,12 - 1,2}{0,21/\sqrt{12}} = -1,32$$

**Langkah 4.** Uji yang dilakukan adalah uji dua arah, taraf signifikansi  $\alpha=0,05$  maka nilai kritis  $t_{\alpha/2}$  dengan derajat kebebasan 11 adalah  $\pm 2,20$ , gambar aturan keputusannya adalah :



**Langkah 5.** Karena  $t = -1,32$  berada di daerah Terima  $H_0$ , maka keputusannya adalah menerima  $H_0$ . Kesimpulannya pada taraf signifikan 0,05, maka akuntan tersebut tidak perlu melaporkan apapun kepada bagian keuangan karena rata-rata invoice penjualan masih sama dengan Rp 1,2 juta.

### 3. Uji Z untuk Pengujian Hipotesis Proporsi

Pada keadaan tertentu, pengujian hipotesis tentang proporsi populasi suatu kategori tertentu lebih diperlukan daripada ukuran rata-rata populasi. Perhatikan pernyataan-pernyataan berikut :

1. Kepala sekolah suatu SMK menyatakan bahwa 80% lulusannya memasuki lapangan kerja yang sesuai dengan bidang ilmu mereka.
2. Seorang manajer pemasaran menyatakan bahwa lebih dari 45% penjualan produk mereka dilakukan melalui internet.
3. Seorang aktivis perempuan menyatakan bahwa keterwakilan wanita di DPR adalah kurang dari 20%.

Seandainya akan dilakukan pengujian untuk pernyataan-pernyataan tersebut, maka uji statistik yang digunakan adalah

$$Z = \frac{p_s - p}{\sqrt{\frac{p(1-p)}{n}}}$$

dimana  $p_s = \text{proporsi sampel} = \frac{x}{n}$  X  $\frac{\text{banyaknya sukses dalam sampel}}{\text{ukuran sampel}}$

$p = \text{proporsi populasi yang dihipotesiskan}$

Langkah-langkah pengujian sama dengan langkah pengujian hipotesis pada bagian bagian sebelumnya.

#### Contoh .

Hasil sebuah survey yang dilakukan oleh perusahaan menunjukkan bahwa 593 dari 1.040 responden memilih untuk diberi uang Rp 500.000 daripada diberi cuti satu hari. Pada taraf signifikan 0,05 , apakah terdapat cukup bukti berdasarkan survey tersebut bahwa lebih dari setengah orang lebih memilih diberi uang Rp 500.000 daripada diberi cuti satu hari?

Jawab :

**Langkah 1.** Kalimat “apakah terdapat cukup bukti berdasarkan survey tersebut bahwa lebih dari setengah orang lebih memilih diberi uang Rp. 500.000 daripada diberi cuti satu hari” menunjukkan bahwa yang menjadi perhatian adalah proporsi

responden yang memilih untuk diberi uang Rp. 500.000. Hipotesis nol adalah proporsi populasi yang memilih untuk diberi uang Rp. 500.000 adalah kurang atau sama dengan 0,5 dan hipotesis alternatif adalah proporsi populasi yang memilih untuk diberi uang Rp. 500.000 adalah lebih dari 0,5. Kedua hipotesis ditulis sebagai berikut :

$$H_0 : p \leq 0,5$$

$$H_1 : p > 0,5$$

Ini merupakan uji satu arah karena hipotesis alternative menyatakan arah perbedaan yaitu *lebih dari*.

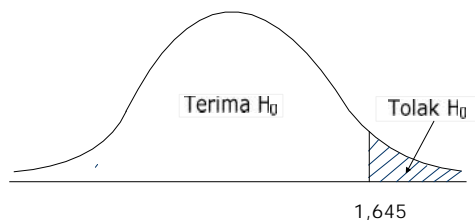
**Langkah 2.** Taraf signifikansi adalah 0,05.

**Langkah 3.** Uji statistik adalah Z dengan

$$Z = \frac{p_s \times \frac{\text{banyaknya responden yang memilih diberi uang Rp 500.000}}{\text{ukuran sampel}} - p_0}{\sqrt{p_0(1-p_0)}} = \frac{0,57 \times \frac{593}{1040} - 0,5}{\sqrt{0,5(1-0,5)}} = 4,51$$

$$Z = \frac{p_s - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0,57 - 0,5}{\sqrt{\frac{0,5(1-0,5)}{1040}}} = 4,51$$

**Langkah 4.** Uji yang dilakukan adalah uji satu arah, taraf signifikansi  $\alpha=0,05$  maka nilai kritis  $Z_{\alpha}=1,645$ , gambar aturan keputusannya adalah :



**Langkah 5.** Karena  $Z=4,51$  berada di daerah Tolak  $H_0$ , maka keputusannya adalah menolak  $H_0$ . Kesimpulannya pada taraf signifikan 0,05, terdapat cukup bukti berdasarkan survey tersebut bahwa lebih dari setengah orang lebih memilih diberi uang Rp 500.000 daripada diberi cuti satu hari.

# MATERI KULIAH: STATISTIKA

<b>Bahan Kajian</b>	<b>:Uji prasyarat (analisa normalitas), homogenitas, dan linearitas</b>
<b>Program Studi</b>	<b>: Pendidikan Teknik Informatika</b>
<b>Minggu ke-</b>	<b>: 7 dan 8</b>

## A. Pembelajaran :

### 1. Learning Outcomes

Memiliki pemahaman yang luas dan teliti tentang uji prasyarat (analisis normalitas), homogenitas, dan linearitas, dan menerapkan menggunakan program SPSS.

### 2. Materi

- a) Normalitas
- b) Homogen dan Linearitas
- c) Pengertian
- d) Penggunaan
- e) Perhitungan

### 3. Uraian Materi

#### A. Normalitas

Uji distribusi normal adalah uji untuk mengukur apakah data yang didapatkan memiliki distribusi normal sehingga dapat dipakai dalam statistik parametrik (statistik inferensial). Dengan kata lain, uji normalitas adalah uji untuk mengetahui apakah data empirik yang didapatkan dari lapangan sesuai dengan distribusi teoritik tertentu atau, distribusi normal atau apakah data yang diperoleh berasal dari populasi yang berdistribusi normal. Penggunaan statistik parametris (data interval/rasio), berkejadengan asumsi bahwa data setiap variabel penelitian yang akan dianalisis membentuk distribusi normal.

Data berdistribusi normal yaitu bahwa data akan mengikuti bentuk distribusi normal, dimana data memusat pada nilai rata-rata dan median. Data yang membentuk distribusi normal bila jumlah data di atas dan di bawah rata-rata adalah sama, demikian juga simpangan bakunya. Data terdistribusi normal adalah salah satu syarat data parametrik sehingga data memiliki karakteristik empirik yang mewakili populasi.



Tujuan uji normalitas adalah ingin mengetahui apakah distribusi sebuah data mengikuti atau mendekati distribusi normal, yakni distribusi data dengan bentuk lonceng (*bell shaped*). Data yang “baik” adalah data yang mempunyai pola seperti distribusi normal, yakni distribusi data tersebut tidak menceng ke kiri atau menceng ke kanan. Uji normalitas bertujuan untuk menguji apakah dalam model regresi, variabel pengganggu atau residual memiliki distribusi normal. Jika asumsi ini dilanggar, maka uji statistik menjadi tidak valid atau bisa terutama untuk sampel kecil.

Uji normalitas dapat dilakukan melalui dua pendekatan yaitu melalui pendekatan grafik (histogram dan P-P Plot) atau Uji Kolmogorov-smirnov, Chi-square, Liliefors maupun Shapiro-Wilk. Ada tiga pilihan yang dapat dilakukan jika diketahui bahwa data tidak normal; yaitu :

1. Jika jumlah sampel besar, maka dapat menghilangkan nilai outlier dari data (bahasan outlier akan dibahas kemudian)
2. Melakukan transformasi data
3. Menggunakan alat analisis nonparametrik

### **1. Uji Liliefors**

Uji normalitas dilakukan sebelum data diolah dengan berbagai teknik seperti: korelasi product moment, regresi, t-test, anava dan sebagainya. Teknik yang sering digunakan untuk uji normalitas data adalah teknik uji Liliefors. Teknik Liliefors menggunakan pendekatan pemeriksaan data individu dalam keseluruhan (kelompok). Data yang diperoleh ditransformasikan dalam nilai Z (yaitu selisih data dengan rata-rata dibandingkan standar deviasi data tersebut). Prosedurnya akan jadi rumit apabila jumlah data cukup banyak. Karena itu, teknik Liliefors biasanya digunakan untuk rentang data yang relatif sedikit. Sedangkan untuk rentangan yang lebih besar digunakan teknik chi kuadrat, dengan menguji data berkelompok. Karena asumsinya normal, maka pengujian didasarkan pada pendekatan Stanine. Hipotesis pengujian sebagai berikut:

Ho: Sampel berasal dari populasi berdistribusi normal.

H1: Sampel berasal dari populasi berdistribusi tidak normal.

Kriteria Pengujian: Tolak Ho, jika  $L_o > L$  kritis, selain itu Ho diterima

Langkah-langkah uji normalitas :

1. Urutkan data sampel dari yang terkecil ke terbesar ( $X_1, X_2, X_3, \dots, X_n$ )
2. Hitung rata-rata nilai skor sampel secara keseluruhan menggunakan rata-rata tunggal.
3. Hitung standar deviasi nilai skor sampel menggunakan standar deviasi tunggal
4. Hitung  $Z_i$  dengan rumus  $Z_i = \frac{X_i - \bar{X}}{s}$
5. Tentukan nilai tabel Z (lihat lampiran tabel z) berdasarkan nilai  $Z_i$ , dengan mengabaikan nilai negatifnya.
6. Tentukan besar peluang masing-masing nilai z berdasarkan tabel z (tuliskan dengan simbol F ( $z_i$ )). Yaitu dengan cara nilai 0,5- nilai tabel Z apabila nilai  $z_i$  negatif (-), dan 0,5 + nilai tabel Z apabila nilai  $z_i$  positif (+)
7. Hitung frekuensi kumulatif nyata dari masing-masing nilai z untuk setiap baris, dan sebut dengan  $S(z_i)$  kemudian dibagi dengan jumlah *number of cases* ( $N$ ) sampel.
8. Tentukan nilai  $L_o$  (hitung) =  $F(z_i) - S(z_i)$  dan bandingkan dengan nilai  $L_{tabel}$  (tabel nilai kritis untuk uji liliefors).  $L_o$  diambil dari nilai yang paling besar dari harga  $F(z_i) - S(z_i)$
9. Apabila  $L_o$  (hitung)  $< L_{tabel}$  maka sampel berasal dari populasi yang berdistribusi normal.

Contoh :

Seorang peneliti ingin mengetahui apakah nilai statistik mahasiswa Tata Boga yang mengambil matakuliah statistik berdistribusi normal? Data dikumpulkan dari 30 orang mahasiswa sebagai berikut :

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
50	54	63	68	71	75	72	69	55	61	64	71	68	75	55

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
72	55	60	69	78	75	77	51	61	65	69	72	60	67	52

Jawab:

1. Tahap pertama cari rata-rata dan standar deviasi data tunggal dengan frekuensi lebih dari satu

2. Rata-rata : 
$$\bar{x} = \frac{\sum fx^2}{f} = \frac{1954}{30} = 65,133$$

3. Standar deviasi : 
$$s = \sqrt{\frac{\sum fx^2}{f} - \bar{x}^2} = \sqrt{\frac{2029,45}{30} - 65,133^2} = 8,224$$

X	f	fx	zi	Tabel z	f(z)	F kum	S (z)	f(z)-S(z)
50	1	50	-1,84	0,4671	0,0329	1	0,03333	0,00043
51	1	51	-1,72	0,4573	0,0427	2	0,06667	0,02397
52	1	52	-1,60	0,4452	0,0548	3	0,10000	0,0452
54	1	54	-1,35	0,4115	0,0885	4	0,13333	0,04483
55	1	55	-1,23	0,3907	0,1093	5	0,16667	0,05737
55	1	55	-1,23	0,3907	0,1093	6	0,20000	0,0907
55	1	55	-1,23	0,3907	0,1093	7	0,23333	0,12403
60	1	60	-0,62	0,2324	0,2676	8	0,26667	0,000933
60	1	60	-0,62	0,2324	0,2676	9	0,30000	0,0324
61	1	61	-0,50	0,1915	0,3085	10	0,33333	0,02483
61	1	61	-0,50	0,1915	0,3085	11	0,36667	0,05817
63	1	63	-0,26	0,1206	0,3794	12	0,40000	0,0206
64	1	64	-0,14	0,0055	0,4945	13	0,43333	0,061167
65	1	65	-0,02	0,008	0,492	14	0,46667	0,025333
67	1	67	0,23	0,091	0,591	15	0,50000	0,091
68	1	68	0,35	0,1368	0,6368	16	0,53333	0,103467

68	1	68	0,35	0,1368	0,6368	17	0,56667	0,070133
69	1	69	0,47	0,1808	0,6808	18	0,60000	0,0808
69	1	69	0,47	0,1808	0,6808	19	0,63333	0,047467
69	1	69	0,47	0,1808	0,6808	20	0,66667	0,014133
71	1	71	0,71	0,4564	0,9564	21	0,70000	0,2564
71	1	71	0,71	0,4564	0,9564	22	0,73333	0,223067
72	1	72	0,83	0,2967	0,7967	23	0,76667	0,030033
72	1	72	0,83	0,2967	0,7967	24	0,80000	0,0033
72	1	72	0,83	0,2967	0,7967	25	0,83333	0,03663
75	1	75	1,20	0,3849	0,8849	26	0,86667	0,018233
75	1	75	1,20	0,3849	0,8849	27	0,90000	0,0151
75	1	75	1,20	0,3849	0,8849	28	0,93333	0,04843
77	1	77	1,44	0,4251	0,9251	29	0,96667	0,04157
78	1	78	1,56	0,4406	0,9406	30	1,00000	0,0594
	30	1954						

Untuk nilai  $X_1$

4. Nilai  $Z_i = \frac{X_i - \bar{X}}{s} = \frac{50 - 65,133}{8,22} = -1,84$

5. Nilai tabel Z (lihat lampiran tabel z) berdasarkan nilai  $Z_i$ , dengan mengabaikan nilai negatifnya yaitu pada kolom pertama 1,8 kemudian lihat pada kolom diatas angka 4 yaitu 0,4671
6. Mentukan besar peluang masing-masing nilai z berdasarkan tabel z (tuliskan dengan simbol  $F(z_i)$ ). Yaitu dengan cara nilai 0,5- nilai tabel Z apabila nilai  $z_i$  negatif (-), dan 0,5 + nilai tabel Z apabila nilai  $z_i$  positif (+). Karena pada  $X_1$  bernilai negatif maka  $F(z_i) = 0,5 - 0,4671 = 0,0329$
7. F kum yaitu frekuensi nya dijumlahkan dengan frekuensi sebelumnya.  
Misalnya untuk  $X_2 = 1 + 1 = 2$
8.  $S(z) =$  frekuensi kumulatif/jumlah frekuensi. Misalnya untuk  $X_1 = 1/30 = 0,033$

9.  $f(z)-S(z) = 0,0329 - 0,0333 = 0,0043$ . Harga yang paling besar dari kolom ini adalah 0,124 ( $L_o$  hitung) sedangkan  $L_o$  tabel = 0,161 (dari tabel khusus untuk uji Liliefors)
10.  $L_o$  hitung <  $L_o$  tabel atau  $0,124 < 0,161$  maka data berdistribusi normal

**Tabel Nilai Kritis Untuk Uji Lilliefors**

Ukuran Sampel	Tarf Nyata ( $\alpha$ )				
	0.01	0.05	0.10	0.15	0.20
n = 4	0.417	0.381	0.352	0.319	0.300
5	0.405	0.337	0.315	0.299	0.285
6	0.364	0.319	0.294	0.277	0.265
7	0.348	0.300	0.276	0.258	0.247
8	0.331	0.285	0.261	0.244	0.233
9	0.311	0.271	0.249	0.233	0.223
10	0.294	0.258	0.239	0.224	0.215
11	0.284	0.249	0.230	0.217	0.206
12	0.275	0.242	0.223	0.212	0.199
13	0.268	0.234	0.214	0.202	0.190
14	0.261	0.227	0.207	0.194	0.183
15	0.257	0.220	0.201	0.187	0.177
16	0.250	0.213	0.195	0.182	0.173
17	0.245	0.206	0.189	0.177	0.169
18	0.239	0.200	0.184	0.173	0.166
19	0.235	0.195	0.179	0.169	0.163
20	0.231	0.190	0.174	0.166	0.160
25	0.200	0.173	0.158	0.147	0.142
30	0.187	0.161	0.144	0.136	0.131
n > 30	<u>1.031</u>	<u>0.886</u>	<u>0.85</u>	<u>0.768</u>	<u>0.736</u>
	$\sqrt{n}$	$\sqrt{n}$	$\sqrt{n}$	$\sqrt{n}$	$\sqrt{n}$

Sumber :

Sudjana, (1992), *Metoda Statistika*, Bandung: Tarsito

## 2. Uji Kolmogorov Smirnov

Banyak sekali teknik pengujian normalitas suatu distribusi data yang telah dikembangkan oleh para ahli. Salah satu pengujian normalitas dengan menggunakan teknik Kolmogorov Smirnov. Uji Kolmogorov Smirnov merupakan pengujian normalitas yang banyak dipakai, terutama setelah adanya banyak program statistik yang beredar. Kelebihan dari uji ini adalah sederhana dan tidak menimbulkan perbedaan persepsi di antara satu pengamat dengan pengamat yang lain, yang sering terjadi pada uji normalitas dengan menggunakan grafik.

Konsep dasar dari uji normalitas Kolmogorov Smirnov adalah dengan membandingkan distribusi data (yang akan diuji normalitasnya) dengan distribusi normal baku. Distribusi normal baku adalah data yang telah ditransformasikan ke dalam bentuk Z-Score dan diasumsikan normal. Jadi sebenarnya uji Kolmogorov Smirnov adalah uji beda antara data yang diuji normalitasnya dengan data normal baku. Seperti pada uji beda biasa, jika signifikansi di bawah 0,05 berarti terdapat perbedaan yang signifikan, dan jika signifikansi di atas 0,05 maka tidak terjadi perbedaan yang signifikan berarti data yang diuji normal, tidak berbeda dengan normal baku..

Kelemahan dari Uji Kolmogorov Smirnov yaitu bahwa jika kesimpulan memberikan hasil yang tidak normal, maka tidak bisa menentukan transformasi seperti apa yang harus digunakan untuk normalisasi. Jadi kalau tidak normal, gunakan plot grafik untuk melihat menceng ke kanan atau ke kiri, atau menggunakan Skewness dan Kurtosis sehingga dapat ditentukan transformasi seperti apa yang paling tepat dipergunakan.

Langkah-langkah dengan metode Kolmogorov smirnov yaitu tidak jauh berbeda dengan metode Liliefors. Dimana penyelesaian dan penggunaan rumus sama, namun pada signifikansi yang berbeda. Signifikansi metode Kolmogorov Smirnov menggunakan tabel pembandingan Kolmogorov Smirnov, sedangkan metode Liliefors menggunakan tabel pembandingan metode Liliefors.

Rumus :

No	$X_i$	$Z_i = \frac{X_i - \bar{X}}{s}$	Fr	Fs	Fr - Fs
1					
2					
3					
dst					

Dimana

$X_i$  = angka pada data

Z = transformasi dari angka ke notasi pada distribusi normal

Fr = Probabilitas kumulatif normal

Fs = Probabilitas kumulatif emperis

Fr = kumulatif proporsi luasan kurva normal berdasarkan notasi  $Z_i$ , dihitung dari luasan kurva mulai dari ujung kiri kurva sampai dengan titik Z

$$F_s = \frac{\text{Banyaknya angka sampai angka ke } n}{\text{Banyaknya seluruh angka pada data}}$$

Persyaratan untuk uji normalitas ini yaitu :

- Data berskala interval atau ratio
- Data tunggal atau belum dikelompokkan pada tabel distribusi frekuensi
- Dapat untuk jumlah data yang besar ataupun kecil

Signifikansi :

Nilai  $Fr - Fs$  terbesar dibandingkan dengan nilai tabel Kormogorov Smarnov.

Jika nilai  $Fr - Fs$  terbesar kurang dari nilai tabel Kormogorov-Smirnov, maka

$H_0$  diterima dan  $H_1$  ditolak. Dan jika  $Fr - Fs$  terbesar dari nilai tabel

Kolmogorov Smirnov, maka  $H_0$  yang ditolak dan  $H_1$  diterima

Tabel Nilai Kritis Uji Kolmogorov-Smirnov

n	$\mathfrak{S} = 0,20$	$\mathfrak{S} = 0,10$	$\mathfrak{S} = 0,05$	$\mathfrak{S} = 0,02$	$\mathfrak{S} = 0,01$
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
6	0,410	0,468	0,519	0,577	0,617
7	0,381	0,436	0,483	0,538	0,576
8	0,359	0,410	0,454	0,507	0,542
9	0,339	0,387	0,430	0,480	0,513
10	0,323	0,369	0,409	0,457	0,486
11	0,308	0,352	0,391	0,437	0,468
12	0,296	0,338	0,375	0,419	0,449
13	0,285	0,325	0,361	0,404	0,432
14	0,275	0,314	0,349	0,390	0,418
15	0,266	0,304	0,338	0,377	0,404
16	0,258	0,295	0,327	0,366	0,392
17	0,250	0,286	0,318	0,355	0,381
18	0,244	0,279	0,309	0,346	0,371
19	0,237	0,271	0,301	0,337	0,361
20	0,232	0,265	0,294	0,329	0,352
21	0,226	0,259	0,287	0,321	0,344
22	0,221	0,253	0,281	0,314	0,337
23	0,216	0,247	0,275	0,307	0,330
24	0,212	0,242	0,269	0,301	0,323
25	0,208	0,238	0,264	0,295	0,317
26	0,204	0,233	0,259	0,290	0,311
27	0,200	0,229	0,254	0,284	0,305
28	0,197	0,225	0,250	0,279	0,300
29	0,193	0,221	0,246	0,275	0,295
30	0,190	0,218	0,242	0,270	0,290
35	0,177	0,202	0,224	0,251	0,269
40	0,165	0,189	0,210	0,235	0,252
45	0,156	0,179	0,198	0,222	0,238
50	0,148	0,170	0,188	0,211	0,226
55	0,142	0,162	0,180	0,201	0,216
60	0,136	0,155	0,172	0,193	0,207
65	0,131	0,149	0,166	0,185	0,199
70	0,126	0,144	0,160	0,179	0,192
75	0,122	0,139	0,154	0,173	0,185
80	0,118	0,135	0,150	0,167	0,179
85	0,114	0,131	0,145	0,162	0,174
90	0,111	0,127	0,141	0,158	0,169
95	0,108	0,124	0,137	0,154	0,165
100	0,106	0,121	0,134	0,150	0,161



Pendekatan 1,07/ n 1,22/ n 1,36/ n 1,52/ n 1,63/ n

### 3. Uji Normalitas dengan $X^2$ (kai kuadrat/chi kuadrat)

Uji normalitas ini digunakan untuk menguji normalitas data dalam bentuk data kelompok dalam distribusi frekuensi.

Langkah-langkah:

- Membuat daftar distribusi frekuensi data kelompok
- Hitung nilai rata-rata data kelompok
- Hitung nilai standar deviasi data kelompok
- Buatlah batas nyata tiap interval kelas dan dijadikan sebagai  $X_i$  ( $X_1, X_2, X_3, \dots, X_n$ ). Nilai  $X_i$  dijadikan bilangan baku  $Z_1, Z_2, Z_3, \dots, Z_n$ . Dimana nilai baku  $Z_i$  ditentukan dengan rumus  $Z_i = \frac{X_i - \bar{X}}{s}$
- Tentukan besar peluang masing-masing nilai z berdasarkan tabel Z, dan sebut dengan F ( $Z_i$ )
- Tentukan luas tiap kelas interval dengan cara mengurangi nilai F ( $Z_i$ ) yang lebih besar di atas atau di bawahnya.
- Tentukan  $f_e$  (frekuensi ekpektasi/frekuensi harapan) dengan cara membagi luas kelas tiap interval dibagi number of cases (N/sampel)
- Masukkan frekuensi absolut sebagai  $f_o$  (frekuensi observasi)
- Cari nilai  $X^2$  tiap interval dengan rumus

$$X^2 = \sum_{i=1}^k \frac{f_o - f_h}{f_h}^2$$

○

- Jumlahkan seluruh  $X^2$  dari keseluruhan kelas interval
- Bandingkan jumlah total  $X^2_{hitung}$  dengan  $X^2_{tabel}$ . Apabila  $X^2_{hitung} < X^2_{tabel}$  maka sampel berasal dari populasi yang berdistribusi normal

## **B. Homogen dan Linearitas**

### **1) HOMOGENITAS**

Pengujian homogenitas dimaksudkan untuk memberikan keyakinan bahwa sekumpulan data yang dimanipulasi dalam serangkaian analisis memang berasal dari populasi yang tidak jauh berbeda keragamannya. Pada dasarnya uji homogenitas dimaksudkan untuk memperlihatkan bahwa dua atau lebih kelompok data sampel berasal dari populasi yang memiliki variansi yang sama. Homogenitas varian merupakan syarat untuk menguji kesamaan dua rata-rata atau lebih. Pengujian homogenitas adalah pengujian mengenai sama tidaknya variansi-variansi dua buah distribusi atau lebih. Jika keseluruhan responden penelitian yang dipilih pada saat penentuan sampel penelitian terdiri dari beberapa unsur (pihak) yang berbeda satu sama lain atau tidak homogen maka hasil pengujian homogenitas harus tidak homogen; dan jika keseluruhan responden penelitian yang dipilih pada saat penentuan sampel penelitian hanya terdiri satu unsur (pihak) saja maka hasil pengujian homogenitas harus homogen.

Apabila hasil pengujian homogenitas tidak sama dengan keadaan keseluruhan responden penelitian (terdiri satu unsur saja, atau terdiri dari beberapa unsur), maka pengolahan data tidak bisa dilanjutkan ke dalam pengukuran pengaruh/hubungan dan pengujian hipotesis. Alasannya, data yang didapatkan dari para responden dianggap tidak merepresentasikan keseluruhan responden secara benar menurut keadaan yang sebenarnya.

Pengujian homogenitas varians suatu kelompok data, dapat dilakukan dengan cara: 1) Uji F dan 2) Uji Bartlett. Adapun proses pengujian dan rumus yang digunakan untuk pengujian homogenitas varians kelompok data yaitu sebagai berikut:

#### **1. Uji F**

Perhitungan homogenitas menggunakan uji F dapat dilakukan dengan cara membandingkan  $F_{\text{tabel}}$  dengan  $F_{\text{hitung}}$ . Jika  $F_{\text{tabel}} < F_{\text{hitung}}$ , maka kelompok data homogen.

Adapun langkah-langkah menghitungnya adalah :

- o hitung nilai varians setiap kelompok data
- o bandingkan nilai varians kedua kelompok data, lalu tentukan nilai varians yang paling besar dan paling kecil
- o F hitung adalah pembagian varians terbesar dengan varians terkecil.

Rumus untuk uji F yaitu :  $F = S_1^2 / S_2^2$

Dimana :  $S_1^2 =$  varians kelompok 1

$S_2^2 =$  varians kelompok 2

Pengujian hipotesis :

$H_0 : \exists_1^2 = \exists_2^2$  (varians data homogen)

$H_1 : \exists_1^2 \neq \exists_2^2$  (varians data tidak homogen)

Kriteria pengujian    Jika: F hitung  $\geq$  F tabel (0,05; dk1; dk2), maka tolak  $H_0$

                                  Jika: F hitung  $<$  F tabel (0,05; dk1; dk2), maka terima  $H_0$

Contoh

Apakah data tentang hubungan antara sarapan (X) dan hasil belajar (Y) yang dipadatkan dari hasil penelitian bersifat homogen. Datanya adalah sebagai berikut:

X	Y
75	68
78	72
38	63
94	74
83	72
91	74
38	58
68	58
565	539

Jawab

No	X	Y	X <sup>2</sup>	Y <sup>2</sup>	XY
1	75	68	5625	4624	5100
2	78	72	6084	5184	5616
3	38	63	1444	3969	2394
4	94	74	8836	5476	6956
5	83	68	6889	4624	5644
6	91	81	8281	6561	7371
7	87	72	7569	5184	6264
8	91	74	8281	5476	6734
9	38	58	1444	3364	2204
10	68	58	4624	3364	3944
Jumlah	743	688	59077	47826	52227

Kemudian dilakukan penghitungan varians dan standar deviasi, dengan rumus yang ada:

$$S_X^2 = \sqrt{\frac{10.59077 - 743^2}{10(10-1)}} = \sqrt{430.23} = 20.74$$

$$S_Y^2 = \sqrt{\frac{10 - 47826 - 688^2}{10(10-1)}} = \sqrt{54.62} = 7.39$$

Kemudian dicari F hitung :

$$F = \frac{S_{\text{besar}}}{S_{\text{kecil}}} = \frac{20.74}{7.39} = 2.81$$

Dari penghitungan diatas diperoleh F hitung 2.81 dan dari grafik daftar distribusi F dengan dk pembilang = 10-1 = 9. Dk penyebut = 10-1 = 9. Dan  $\alpha = 0.05$  dan F

tabel = 3.18. Tampak bahwa F hitung < [F tabel](#). Hal ini berarti data variabel X dan Y homogen.

df untuk penyebut (N2)	df untuk pembilang (N1)														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	161	199	216	225	230	234	237	239	241	242	243	244	245	245	246
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41	19.42	19.42	19.43
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85

## 2. Uji Barlett

Uji Bartlett digunakan untuk menguji apakah sejumlah sampel berasal dari populasi dengan varians yang sama. Sejumlah sampel bisa berapa saja, karena biasanya uji Bartlett digunakan untuk menguji sampel atau kelompok yang lebih dari 2. Varians yang sama di seluruh sampel disebut homogenitas varians. Uji Bartlett pertama kali diperkenalkan oleh M. S. Bartlett (1937). Uji Bartlett diperlukan dalam beberapa uji statistik seperti [analysis of variance \(ANOVA\)](#).

Perhitungan homogenitas menggunakan metoda uji Bartlett menggunakan rumusan berikut :

- Varians gabungan dari semua sampel:

$$s^2 = \frac{\sum (n_i - 1) s_i^2}{\sum (n - 1)}$$

- Harga satuan B dengan rumus:

$$B = (\log s^2) \sum (n_i - 1)$$

- Uji Bartlett digunakan statistik chi-kuadrat yaitu :

$$\chi^2 = (\ln 10) \left\{ B - \sum (n-1) \log s_i^2 \right\}$$

Dengan  $\ln 10 = 2.3026$ .

**Signifikansi:**

Jika  $\chi^2 \geq \chi_{(1-\alpha)(k-1)}^2$  maka  $H_0$  ditolak

Dimana jika  $\chi^2 \geq \chi_{(1-\alpha)(k-1)}^2$  didapatkan dari tabel distribusi Chi-Kuadrat dengan peluang  $(1- \alpha)$  dan  $df = (k-1)$

Setelah nilai Chi-Kuadrat hitung diperoleh, maka nilai Chi-Kuadrat tersebut dibandingkan dengan Chi-Kuadrat tabel. Kriteria homogen ditentukan *jika* Chi-Kuadrat hitung  $<$  Chi Kuadrat tabel.

Tabel Distribusi  $\chi^2$ 

$\alpha$	0.1	0.05	0.025	0.01	0.005	
db	1	2.70554	3.84146	5.02390	6.63489	7.87940
	2	4.60518	5.99148	7.37778	9.21035	10.59653
	3	6.25139	7.81472	9.34840	11.34488	12.83807
	4	7.77943	9.48773	11.14326	13.27670	14.86017
	5	9.23635	11.07048	12.83249	15.08632	16.74965
	6	10.64464	12.59158	14.44935	16.81187	18.54751
	7	12.01703	14.06713	16.01277	18.47532	20.27774
	8	13.36156	15.50731	17.53454	20.09016	21.95486
	9	14.68366	16.91896	19.02278	21.66605	23.58927
	10	15.98717	18.30703	20.48320	23.20929	25.18805
	11	17.27501	19.67515	21.92002	24.72502	26.75686
	12	18.54934	21.02606	23.33666	26.21696	28.29966
	13	19.81193	22.36203	24.73558	27.68818	29.81932
	14	21.06414	23.68478	26.11893	29.14116	31.31943
	15	22.30712	24.99580	27.48836	30.57795	32.80149
	16	23.54182	26.29622	28.84532	31.99986	34.26705
	17	24.76903	27.58710	30.19098	33.40872	35.71838
	18	25.98942	28.86932	31.52641	34.80524	37.15639
	19	27.20356	30.14351	32.85234	36.19077	38.58212
	20	28.41197	31.41042	34.16958	37.56627	39.99686
	21	29.61509	32.67056	35.47886	38.93223	41.40094
	22	30.81329	33.92446	36.78068	40.28945	42.79566
	23	32.00689	35.17246	38.07561	41.63833	44.18139
	24	33.19624	36.41503	39.36406	42.97978	45.55836
	25	34.38158	37.65249	40.64650	44.31401	46.92797
	26	35.56316	38.88513	41.92314	45.64164	48.28978
	27	36.74123	40.11327	43.19452	46.96284	49.64504
	28	37.91591	41.33715	44.46079	48.27817	50.99356
	29	39.08748	42.55695	45.72228	49.58783	52.33550
	30	40.25602	43.77295	46.97922	50.89218	53.67187

# MATERI KULIAH: STATISTIKA

**Bahan Kajian** : Uji Korelasi  
**Program Studi** : Pendidikan Teknik Informatika  
**Minggu ke-** : 10 dan 11

## A. Pembelajaran :

### 1. Learning Outcomes

Memiliki kemampuan dalam mengolah data dengan uji korelasi

### 2. Materi

- a) Spearman
- b) Pearson.
- c) Parsial.

### 3. Uraian Materi

Korelasi adalah derajat hubungan antara variabel-variabel statistik yang mengandung tingkat hubungan atau kerjasama diantara dua variabel. Korelasi dapat juga diistilahkan yang digunakan untuk mengukur kekuatan hubungan antara variabel. Analisis korelasional merupakan cara untuk mengetahui ada tidaknya hubungan antar 2 variabel. Jika terjadi hubungan antar variabel maka perubahan yang terjadi pada satu variabel akan menyebabkan terjadinya perubahan pada variabel yang lainnya.

1. Pearson correlation adalah statistik bivariat yang mengandung tingkat hubungan linear diantara dua variabel kuantitatif. Korelasi mengukur derajat hubungan antara dua atau lebih variabel. Hubungan antara dua variabel (Misal X dan Y) dapat linear, non-linear, positif atau negatif. Dua variabel dikatakan berhubungan jika perilaku variabel yang satu mempengaruhi variabel yang lain. Jika tidak terjadi pengaruh, maka kedua variabel tersebut disebut independen. Dalam melakukan analisis korelasi langkah awal yang perlu dilakukan adalah mengidentifikasi variabel yang hendak dikorelasikan. Identifikasi dilakukan untuk menentukan posisi variabel apakah sebagai variabel X (pengaruh) dan variabel (terpengaruh).

Metode analisis korelasi terhadap data, tidak hanya yang terdiri dari satu karakteristik saja. Banyak persoalan atau fenomena yang



meliputi lebih dari sebuah variabel, misalnya berat orang dewasa sampai taraf tertentu bergantung pada tinggi badannya, keterampilan melempar ke target bergantung pada koordinasi mata-tangan, hasil produksi padi bergantung jumlah pupuk. Dalam korelasi juga perlu mempelajari analisis data yang terdiri dari banyak variabel dan bagaimana mereka berhubungan. Analisis korelasi merupakan studi yang membahas tentang derajat (seberapa kuat) hubungan antara dua variabel atau lebih.

Ukuran derajat hubungan disebut koefisien korelasi. Koefisien korelasi merupakan angka yang menunjukkan *arah* dan *kuat* hubungan antara dua variabel atau lebih. Bentuk atau jenis hubungan ada 4 ; yaitu; (a) korelasi positif; (b) korelasi negatif; (c). tidak ada korelasi; (d). korelasi sempurna.

- a) Arah hubungan positif, berarti apabila nilai variabel ditingkatkan , maka akan meningkatkan nilai variabel yang lain. Apabila nilai variabel diturunkan , maka akan menurunkan nilai variabel yang lain.
- b) Arah hubungan negatif, berarti apabila nilai variabel ditingkatkan , maka akan menurunkan nilai variabel yang lain. Apabila nilai variabel diturunkan , maka akan meningkatkan nilai variabel yang lain.
- c) Tidak ada korelasi : jika kedua variabel tidak memperlihatkan adanya hubungan. Ketika variabel A naik, variabel lain B naik tapi pada saat bersamaan variabel B juga bisa turun.
- d) Korelasi sempurna adalah jenis korelasi di mana kenaikan atau penurunan variabel A berbanding dengan kenaikan atau penurunan variabel B.

Uji korelasi atau analisa hubungan banyak digunakan dalam kehidupan sehari-hari. Di dunia ini tidak ada yang dapat hidup sendiri, tetapi memerlukan hubungan dengan orang lain. Hubungan itu pada

umumnya dilakukan dengan maksud tertentu seperti mendapat keringanan pajak, memperoleh kredit, meminjam uang, serta minta pertolongan/bantuan lainnya. Seperti diketahui, pada semua kejadian, baik kejadian ekonomi maupun lainnya, pasti ada faktor yang menyebabkan terjadinya kejadian-kejadian tersebut misalnya merosotnya hasil penjualan pangan mungkin disebabkan karena kalah bersaing dengan pangan impor atau lebih mahal dari pangan impor, merosotnya produksi padi mungkin karena pupuknya berkurang, kurangnya air dan lain sebagainya.

Uraian di atas menunjukkan adanya hubungan (korelasi) antara kejadian yang satu dengan kejadian lainnya. Kejadian itu dapat dinyatakan dengan perubahan nilai variabel. Hubungan antara dua kejadian dapat dinyatakan dengan hubungan dua variabel. Di dalam bab ini kita hanya membahas hubungan linear antara dua variabel X dan Y. Apabila dua variabel X dan Y mempunyai hubungan, maka nilai variabel X yang sudah diketahui dapat dipergunakan untuk memperkirakan/menaksir Y. Peramalan pada dasarnya merupakan perkiraan/ taksiran mengenai terjadinya suatu kejadian.

Variabel Y yang nilainya akan diramalkan disebut variabel tidak bebas, sedangkan variabel X yang nilainya dipergunakan untuk meramalkan nilai Y disebut variabel bebas atau variabel peramal dan seringkali disebut variabel yang menerangkan. Jadi, jelas analisis korelasi ini memungkinkan kita untuk mengetahui sesuatu di luar hasil penyelidikan.

#### Tipe korelasi

- Pearson Product Moment Correlation
  - Sangat umum (diasumsikan bahwa korelasi adalah Pearson  $r$  kecuali kalau spesifikasi sebaliknya)
  - Hubungan Linear
  - Hanya untuk skala Interval atau Ratio

- Spearman Correlation
  - Skala Ordinal
  - Mengandung konsistensi terlepas dari bentuk hubungan

#### A. Pearson Product Momen Correlation

Korelasi Pearson Product Moment ( $r$ ) merupakan korelasi yang sering dipakai oleh mahasiswa dan peneliti. Korelasi ini dikemukakan oleh Karl Pearson tahun 1900. Kegunaannya untuk mengetahui derajat hubungan dan kontribusi variabel bebas (independen) dengan variabel terikat (dependent).

Teknik analisis Korelasi Pearson termasuk teknik statistik para metrik yang menggunakan interval dan ratio dengan persyaratan tertentu. Persyaratan korelasi Pearson antara lain: data dipilih secara acak (random); datanya berdistribusi normal; data yang dihubungkan berpola linier; dan data yang dihubungkan mempunyai pasangan yang sama sesuai dengan subjek yang sama. Kalau salah satu tidak terpenuhi persyaratan tersebut analisis korelasi tidak dapat dilakukan. Rumus yang digunakan pada korelasi Pearson adalah:

$$r_{xy} = \frac{n \sum XY - \sum X \sum Y}{\sqrt{\{n \sum X^2 - (\sum X)^2\} \{n \sum Y^2 - (\sum Y)^2\}}}$$

Dimana :  $\sum XY$  = jumlah perkalian X dan Y

$\sum X^2$  = jumlah kuadrat X

$\sum Y^2$  = jumlah kuadrat Y

n = banyak pasangan nilai

Korelasi Pearson dilambangkan ( $r$ ) dengan ketentuan nilai  $r$  tidak lebih dari harga  $(-1 < r < + 1)$ . Apabila nilai  $r = -1$  artinya korelasinya negatif sempurna, setiap individu mengandung nilai  $z$  yang sama persis pada kedua variabel tetapi dengan tanda yang berkebalikan.;  $r = 0$  artinya tidak ada korelasi dan  $r = 1$  berarti korelasinya sangat kuat, setiap individu mengandung nilai  $z$  yang sama persis pada kedua variabel. Nilai  $r$  tinggi (mendekati 1 atau  $-1$ ) mengindikasikan hubungan yang lebih erat. Nilai  $r$  rendah (mendekati 0)

mengindikasikan hubungan yang lebih lemah. Hubungan yang mendekati 0 mengindikasikan hubungan yang tidak linear sehingga perubahan X tidak cocok untuk memprediksi perubahan variabel Y.

$$r = \frac{\sum Z_x Z_y}{n}$$

Sedangkan arti harga r akan dikonsultasikan dengan tabel interpretasi nilai r sebagai berikut.

Interpretasi Koefisien Korelasi Nilai r

Interval Koefisien	Tingkat Hubungan
0,80 – 1,000	Sangat Kuat
0,60 – 0,799	Kuat
0,40 – 0,599	Cukup Kuat
0,20 – 0,399	Rendah
0,00 – 0,199	Sangat Rendah

Selanjutnya untuk menyatakan besar kecilnya sumbangan variabel X terhadap Y dapat ditentukan dengan rumus koefisien diterminan sebagai berikut.

$$KP = r^2 \times 100\%$$

keterangan: KP = Nilai Koefisien Diterminan  
r = Nilai Koefisien Korelasi

Koefisien determinasi ( $r^2$ ) adalah menentukan seberapa besar nilai X dapat menjelaskan nilai Y atau seberapa besar nilai X dapat mempengaruhi nilai Y sedangkan koefisien korelasi (r) merupakan keeratatan hubungan antara variabel X dengan Y.

Walaupun telah diperoleh nilai koefisien korelasi dari hasil perhitungan di atas, namun keberartian nilai tersebut perlu di uji secara statistik.

Hipotesis yang diuji adalah :

Ho : Koefisien korelasi adalah sama dengan nol

Ha : Koefisien korelasi tidak sama dengan nol, atau berarti

Kriteria pengujiannya :

$H_0$  ditolak jika nilai t-hitung lebih besar daripada t-tabel dengan derajat bebas  $n-2$ , dan demikian pula sebaliknya.

Pengujian lanjutan yaitu uji signifikansi yang berfungsi apabila peneliti ingin mencari makna hubungan variabel X terhadap Y, maka hasil korelasi Pearson tersebut diuji dengan uji Signifikansi dengan rumus :

$$t_{\text{hitung}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

keterangan:  $t_{\text{hitung}}$  = Nilai t

r = Nilai Koefisien korelasi

n = Jumlah Sampel

Contoh : hubungan motivasi dengan hasil kinerja dosen Boga UNP

Motivasi (X) : 60; 70; 75; 65; 70; 60; 80; 75; 85; 90; 70; dan 85

Kinerja (Y) : 450; 475; 450; 470; 475; 455; 475; 470; 485; 480; 475; dan 480.

Pertanyaan ;

- Berapakah besar hubungan motivasi dengan hasil kinerja dosen ?
- Berapakah besar sumbangan (kontribusi) motivasi dengan kinerja dosen?
- Buktikan apakah ada hubungan yang signifikan motivasi dengan kinerja dosen?

**Langkah-langkah menjawab:**

**Langkah 1.**

Membuat  $H_a$  dan  $H_0$  dalam bentuk kalimat :

$H_a$  : ada hubungan yang signifikan motivasi dengan kinerja dosen.

$H_0$  : Tidak ada hubungan yang signifikan motivasi dengan kinerja dosen.

**Langkah 2.**

Membuat  $H_a$  dan  $H_0$  dalam bentuk statistik;

$H_a$  :  $r \neq 0$

$H_0$  :  $r = 0$

**Langkah 3.**

Membuat tabel penolong untuk menghitung korelasi Pearson:

No	X	Y	X <sup>2</sup>	Y <sup>2</sup>	XY
1.	60	450	3600	202500	27000
2.	70	475	4900	225625	33250
3.	75	450	5625	202500	33750
4.	65	470	4225	220900	30550
5.	70	475	4900	225625	33250
6.	60	455	3600	207025	27300
7.	80	475	6400	225625	38000
8.	75	470	5625	220900	35250
9.	85	485	7225	235225	41225
10.	90	480	8100	230400	43200
11.	70	475	4900	225625	33250
12.	85	480	7225	230400	40800
Jumlah	885	5640	66325	2652350	416825

**Langkah 4.**

Mencari  $r_{hitung}$  dengan cara masukkan angka statistik dari tabel penolong dengan rumus ;

$$r_{xy} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{\{n \sum X^2 - (\sum X)^2\} \{n \sum Y^2 - (\sum Y)^2\}}}$$

$$r_{xy} = \frac{12(416.825) - (885)(5.460)}{\sqrt{\{12.(66.325) - (885)^2\} \cdot \{12.(2.652.350) - (5.640)^2\}}}$$

$$r_{xy} = \frac{169.900}{\sqrt{133.463.835.000}} = \frac{169.00}{365.327,02} = 0,465$$

Interpretasi nilai koefisien korelasi cukup kuat atau dalam range 0,4 – 0,599

### Langkah 5.

Mencari besarnya sumbangan (kontribusi) variabel X terhadap Y dengan

rumus :

$$KP = r^2 \times 100\% = 0,465^2 \times 100\% = 21,62 \%$$

Artinya motivasi memberikan kontribusi terhadap kinerja dosen sebesar 21,62% dan sisanya 78,38% ditentukan oleh variabel lain.

### Langkah 6.

Menguji signifikan dengan rumus  $t_{hitung}$  :

$$t_{hitung} = \frac{r\sqrt{n}}{\sqrt{1-r^2}} = \frac{0,465\sqrt{12}}{\sqrt{1-0,465^2}} = \frac{1,55}{0,88} = 1,76$$

Kaidah pengujian :

Jika  $t_{hitung} > t_{tabel}$ , maka tolak  $H_0$  artinya signifikan dan

$t_{hitung} < t_{tabel}$ , terima  $H_0$  artinya tidak signifikan.

Berdasarkan perhitungan di atas ,  $\alpha = 0,05$  dan  $n = 12$ , uji dua pihak;  $dk = n - 2 = 12 - 2 = 10$  sehingga diperoleh  $t_{tabel} = 2,228$ . Ternyata  $t_{hitung}$  lebih besar dari  $t_{tabel}$ , atau  $1,76 > 2,228$ , maka  $H_0$  ditolak, artinya ada hubungan yang signifikan motivasi dengan kinerja dosen.

### Langkah 7.

Membuat kesimpulan

1. Hubungan motivasi dengan kinerja dosen adalah  $r_{xy}$  sebesar 0,465 kategori cukup kuat.
2. Kontribusi motivasi dengan kinerja dosen sebesar  $KP = r^2 \times 100\% = 0,465^2 \times 100\% = 21,62\%$ . Artinya motivasi memberikan kontribusi terhadap kinerja dosen sebesar 21,62% dan sisanya 78,38% ditentukan oleh variable lain.

3. Dan terbukti bahwa ada hubungan yang signifikan motivasi dengan kinerja dosen, karena  $t_{hitung}$  lebih besar dari  $t_{tabel}$ , atau  $3,329 > 2,228$ , maka  $H_0$  ditolak, artinya ada hubungan yang signifikan motivasi dengan kinerja dosen.

Contoh lain yaitu hubungan jumlah pengunjung restoran dengan keuntungan.

Jumlah pengunjung dinyatakan dengan  $x$  dan keuntungan dinyatakan dengan  $y$  sebagai berikut:

x	y
3	8
2	6
2	7
1	5
4	6
3	7
1	5
4	9
4	6
2	6

**Langkah 1.**

Membuat  $H_a$  dan  $H_0$  dalam bentuk kalimat :

$H_a$  : ada hubungan yang signifikan jumlah pengunjung dengan keuntungan.

$H_0$  : Tidak ada hubungan yang signifikan jumlah pengunjung dengan keuntungan.

**Langkah 2.**

Membuat  $H_a$  dan  $H_0$  dalam bentuk statistik;

$H_a$  :  $r \neq 0$

$H_0$  :  $r = 0$

**Langkah 3.**

Membuat tabel penolong untuk menghitung korelasi Pearson:



No	x	y	X <sup>2</sup>	Y <sup>2</sup>	XY
1	3	8	9	64	24
2	2	6	4	36	12
3	2	7	4	49	14
4	1	5	1	25	5
5	4	6	16	36	24
6	3	7	9	49	21
7	1	5	1	25	5
8	4	9	16	81	36
9	4	6	16	36	24
10	2	6	4	36	12
	26	65	80	437	177

$$r_{xy} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{\{n \sum X^2 - (\sum X)^2\} \{n \sum Y^2 - (\sum Y)^2\}}}$$

$$r_{xy} = \frac{10(177) - (26)(65)}{\sqrt{\{10(80) - (26)^2\} \cdot \{10(437) - (65)^2\}}}$$

$$r_{xy} = \frac{80}{\sqrt{17980}} = \frac{80}{134,0895} = 0,59661$$

Interpretasi nilai koefisien korelasi cukup kuat atau dalam range 0,4 – 0,599

#### Langkah 5.

Mencari besarnya sumbangan (kontribusi) variabel X terhadap Y dengan

rumus :

$$KP = r^2 \times 100\% = 0,59661^2 \times 100\% = 35,59 \%$$

Artinya jumlah pengunjung memberikan kontribusi terhadap keuntungan sebesar 35,595% dan sisanya 64,41% ditentukan oleh variabel lain.

#### Langkah 6.

Menguji signifikan dengan rumus  $t_{hitung}$  :

$$t_{hitung} = \frac{r\sqrt{nZ^2}}{\sqrt{1-Zr^2}} = \frac{0,59661\sqrt{10Z^2}}{\sqrt{1-0,59661^2}} = \frac{1,687}{0,6351} = 2,657$$

Kaidah pengujian :

Jika  $t_{hitung} > t_{tabel}$ , maka tolak  $H_0$  artinya signifikan dan  
 $t_{hitung} < t_{tabel}$ , terima  $H_0$  artinya tidak signifikan.

Berdasarkan perhitungan di atas ,  $\alpha = 0,05$  dan  $n = 10$ , uji dua pihak;  $dk = n - 2 = 10 - 2 = 8$  sehingga diperoleh  $t_{tabel} = 2,306$ . Ternyata  $t_{hitung}$  lebih besar dari  $t_{tabel}$ , atau  $2,657 > 2,306$ , maka  $H_0$  ditolak, artinya ada hubungan yang signifikan jumlah pengunjung restoran dengan keuntungannya.

## B. Korelasi Spearman

Korelasi Spearman merupakan statistik yang didasarkan atas rangking atau jenjang. Koefisien korelasi rank Spearman adalah yang paling awal dikembangkan dan mungkin yang dikenal baik hingga kini. Statistik ini sering juga disebut dengan rho, disini ditulis dengan  $r_s$ , ini adalah ukuran hubungan yang menuntut kedua variabel diukur sekurang-kurangnya dalam skala ordinal sehingga objek yang dipelajari dapat di-rangking atau dua rangkaian berurut.

Koefisien korelasi Spearman merupakan statistik nonparametrik. Korelasi Spearman digunakan jika data variabel ordinal (berjenjang atau peringkat) yang disebut juga korelasi nonparametrik. Korelasi Spearman merupakan suatu ukuran hubungan yang dapat digunakan pada kondisi satu atau kedua variabel adalah kuantitatif namun kondisi normal tidak terpenuhi. Simbol ukuran sampelnya adalah  $r_s$ . Rumus untuk korelasi Spearman adalah:

$$r_s = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N^3 - N}$$

dimana :  $d_i$  adalah perbedaan antara kedua rangking  
 $N$  adalah banyaknya observasi

Pembuatan ranking dapat dimulai dari nilai terkecil atau nilai terbesar tergantung permasalahannya. Bila ada data yang nilainya sama, maka pembuatan rangking didasarkan pada nilai rata-rata dari rangking-rangking data tersebut. Jika proporsi

angka yang sama tidak besar, maka rumus diatas masih bisa digunakan. Namun jika proporsi angka yang sama cukup besar, maka dapat digunakan suatu faktor koreksi dan rumus menjadi :

$$r_s = \frac{2\left(\frac{N^3 - N}{12}\right) - \sum T_1 - \sum T_2 - \sum d_i^2}{2\sqrt{\left(\frac{N^3 - N}{12} - \sum T_1\right)\left(\frac{N^3 - N}{12} - \sum T_2\right)}} \quad \text{dimana:} \quad T = \frac{t^3 - t}{12}$$

$t$  adalah banyaknya observasi yang berangka sama pada suatu ranking tertentu.

Contoh seorang manager produksi ingin mengetahui apakah ada hubungan antara nilai tes bakat pada waktu penerimaan kerja dengan *rating* tampilannya setelah satu semester bekerja. Nilai tes bakat berkisar antara 0 sampai 100, sedangkan *rating* tampilan mempunyai skala sebagai berikut:

- 1 = pekerja berpenampilan sangat dibawah rata-rata
- 2 = pekerja berpenampilan dibawah rata-rata
- 3 = pekerja berpenampilan sedang (rata-rata)
- 4 = pekerja berpenampilan diatas rata-rata
- 5 = pekerja berpenampilan sangat diatas rata-rata

Pekerja	nilai test	rating	Pekerja	nilai test	rating
1	59	3	11	48	3
2	47	2	12	65	3
3	58	4	13	51	2
4	66	3	14	61	3
5	77	2	15	40	3
6	57	4	16	67	4
7	62	3	17	60	2
8	68	3	18	56	3
9	69	5	19	76	3
10	36	1	20	71	5

Jawab

Pekerja	nilai test	rank(a)	rating	rank(b)	$d_i$	$d_i^2$
1	59	9	3	10,5	-1,5	2,25
2	47	3	2	3,5	-0,5	0,25
3	58	8	4	17	-9	81
4	66	14	3	10,5	3,5	12,25
5	77	20	2	3,5	16,5	272,25
6	57	7	4	17	-10	100
7	62	12	3	10,5	1,5	2,25
8	68	16	3	10,5	5,5	30,25
9	69	17	5	19,5	-2,5	6,25
10	36	1	1	1	0	0
11	48	4	3	10,5	-6,5	42,25
12	65	13	3	10,5	2,5	6,25
13	51	5	2	3,5	1,5	2,25
14	61	11	3	10,5	0,5	0,25
15	40	2	3	10,5	-8,5	72,25
16	67	15	4	17	-2	4
17	60	10	2	3,5	6,5	42,25
18	56	6	3	10,5	-4,5	20,25
19	76	19	3	10,5	8,5	72,25
20	71	18	5	19,5	-1,5	2,25
					$\Sigma$	771

$$r_s = 1 - \frac{6(771)}{20^3 - 20} = 0,4203$$

Jika menggunakan faktor koreksi  $\sum T_1 = 0$  karena tidak ada nilai yang sama pada variabel nilai test, sedangkan untuk variabel rating  $\sum T_2 = 90$  (nilai 2 ada 4, nilai 3 ada 10, nilai 4 ada 3, dan nilai 5 ada 2). Sehingga nilai  $r_s$  adalah sebesar:

$$r_s = \frac{2\left(\frac{20^3 - 20}{12}\right) - 0 - 90 - 771}{2\sqrt{\left(\frac{20^3 - 20}{12} - 0\right)\left(\frac{20^3 - 20}{12} - 90\right)}} = 0,3792.$$

Uji untuk menentukan adanya hubungan antara dua variabel adalah:

Untuk sampel kecil ( $n < 30$ ) → gunakan tabel koefisien korelasi Spearman.

Keputusan : tolak  $H_0$  jika  $r_s < -r_{\text{tabel}}$  dan  $r_s > r_{\text{tabel}}$  (pada  $n$  dan  $\alpha$  tertentu). Untuk sampel besar ( $n > 30$ ) dapat didekati dengan distribusi normal.

$$z = \frac{r_s - 0}{1/\sqrt{n-1}} = r_s \sqrt{n-1}$$

Daerah tolak  $H_0$  :  $Z < -Z_{\alpha/2}$  atau  $Z > Z_{\alpha/2}$

# MATERI KULIAH: STATISTIKA

**Bahan Kajian** : Uji t  
**Program Studi** : Pendidikan Teknik Informatika  
**Minggu ke-** : 13 dan 14

## A. Pembelajaran :

### 1. Learning Outcomes

Memiliki kemampuan tentang uji t

### 2. Materi

- a) Uji t satu sampel bebas
- b) Uji t dua sampel berpasangan
- c) Uji t dua sampel bebas

### 3. Uraian Materi

#### A. Uji t Satu Sampel

#### B. Uji t Dua Sampel

Uji-t atau t-test merupakan statistik uji yang sering kali ditemui dalam masalah-masalah praktis statistika. Uji t pertama kali dikembangkan oleh William Seely Gosset pada 1915. Awalnya ia menggunakan nama samaran *Student*, dan huruf *t* yang terdapat dalam istilah Uji “t “ dari huruf terakhir nama beliau. Uji t disebut juga dengan nama *Student t*. Uji-t termasuk dalam golongan statistika parametrik. Statistik uji ini digunakan dalam pengujian hipotesis, uji-t digunakan ketika informasi mengenai nilai variance (ragam) populasi tidak diketahui. Uji-t adalah salah satu uji yang digunakan untuk mengetahui ada atau tidaknya perbedaan yang signifikan dari satu atau dua rata-rata sampel (dua buah variabel yang dikomparasikan). Persyaratan analisa uji t yaitu sampel di ambil secara acak dari populasi berdistribusi normal dan data berskala interval dan atau rasio. Sedangkan uji t berguna sebagai alat analisis data untuk menguji satu sampel atau dua sampel dan membandingkan dua rata-rata untuk menentukan apakah perbedaan rata-rata tersebut perbedaan nyata atau karena kebetulan.

Uji-t dapat dibagi menjadi 2, yaitu uji-t yang digunakan untuk pengujian hipotesis 1-sampel dan uji-t yang digunakan untuk pengujian hipotesis 2-sampel. Bila dihubungkan dengan kebebasan (independen) sampel yang

digunakan (khusus bagi uji-t dengan 2-sampel), maka uji-t dibagi lagi menjadi 2, yaitu uji-t untuk sampel bebas (independent) dan uji-t untuk sampel berpasangan (paired). Ada tiga bentuk hipotesis untuk uji t dimana penggunaannya tergantung dari persoalan yang akan diuji yaitu:

- a. Bentuk uji hipotesis dua sisi dengan hipotesis  $H_0 \mu = \mu_0$  dan  $H_1 \mu \neq \mu_0$
- b. Bentuk uji hipotesis satu sisi untuk sisi atas  $H_0 \mu \leq \mu_0$  dan  $H_1 \mu > \mu_0$
- c. Bentuk uji hipotesis dengan satu sisi untuk sisi bawah  $H_0 \mu \geq \mu_0$  dan  $H_1 \mu < \mu_0$

Dalam pengujian hipotesis, kriteria untuk menolak atau tidak menolak  $H_0$  berdasarkan p-value adalah sebagai berikut:

Jika p-value  $< \alpha$ , maka  $H_0$  ditolak dan jika p-value  $> \alpha$  maka  $H_0$  diterima

Dalam lingkup uji-t untuk pengujian hipotesis 2-sampel bebas, maka ada 1 hal yang perlu mendapat perhatian, yaitu apakah ragam populasi diasumsikan homogen (sama) atau tidak. Bila ragam populasi diasumsikan sama, maka uji-t yang digunakan adalah uji-t dengan asumsi ragam homogen, sedangkan bila ragam populasi dari 2-sampel tersebut tidak diasumsikan homogen, maka yang lebih tepat adalah menggunakan uji-t dengan asumsi ragam tidak homogen. Uji-t dengan ragam homogen dan tidak homogen memiliki rumus hitung yang berbeda. Oleh karena itulah, apabila uji-t hendak digunakan untuk melakukan pengujian hipotesis terhadap 2-sampel, maka harus dilakukan pengujian mengenai asumsi *kehomogenan* ragam populasi terlebih dahulu dengan menggunakan uji-F.

#### A. UJI t SATU SAMPEL

Uji t untuk menguji apakah satu sampel sama atau berbeda dengan rata-rata populasinya. Uji t pada prinsipnya menguji apakah suatu nilai tertentu (yang diberikan sebagai pembanding) berbeda secara nyata ataukah tidak dengan rata-rata sebuah sampel. Nilai yang dimaksud pada umumnya adalah nilai parameter untuk mengukur suatu populasi

Beberapa contoh yang berhubungan dengan uji t satu sampel :

Diduga rata-rata konsumsi mi pada anak kos prodi tata boga 2014 UNP adalah 3 bungkus/minggu. Jika anak kos prodi tata boga 2014 dianggap

populasi maka angka 3 merupakan nilai parameter. Kemudian diambil beberapa sample dan dihitung rata-ratanya. Uji t satu sampel digunakan untuk membandingkan nilai parameter dengan nilai rata-rata dari sample.

Langkah-langkahnya :

- hitung rata-rata dan standar deviasi (s)
- Hitung drajat bebas  $df = n - 1$
- Tentukan tingkat signifikansi ( $\alpha = 0.025$  atau  $0.05$ )
- Pengujiannya apakah menggunakan 1 ekor atau 2 ekor
- Setelah diperoleh t hitung ; lalu bandingkan dengan t tabel : jika t hitung > t tabel  $H_0$  ditolak

Rumus uji t dengan satu sampel :

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

Dimana:

$t$  = Nilai t hitung

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

$\mu$  = nilai parameter

$S$  = Standar deviasi sample

$n$  = Jumlah sample

Contoh: seorang peneliti ingin melakukan kajian tentang kemampuan ujian peserta untuk masuk ke jurusan IKK UNP. Untuk keperluan penelitian ini di ambil sampel sebanyak 49 peserta, yang dipilih secara acak. Standar kelulusan ujian adalah skor 60 (rata-rata populasi). Dari sampel diperoleh rata-rata skor ujian adalah 55 dengan standar deviasi 15. Berdasarkan data ini , pihak UNP membuat pernyataan: Semua peserta ujian mempunyai kemampuan menyelesaikan soal ujian di bawah standar kelulusan.”

Pertanyaan: Ujilah pernyataan (hipotesis di atas),

Jawab :  $H_0 : \mu \geq 60$

$H_1 : \mu < 60$

$\bar{x}$  X55

$\mu$  X60

$S$  X15

$n$  X49

$$t = \frac{55 - 60}{15 / \sqrt{49}} = -2,14 < -2,33$$



Bandingkan hasil perhitungan t di atas dengan tabel t.

$$db = n - 1 = 49 - 1 = 48$$

$$t_{\text{tabel}} = t(r) \text{ X1,679.}$$

$$t > t_{\text{tabel.}} \quad H_0 \text{ ditolak.}$$

Kesimpulan : terdapat peserta ujian mempunyai kemampuan di bawah rata-rata.

Latihan :

Terdapat asumsi bahwa dari suatu populasi mahasiswa Tata boga semester 3 rata-rata nilai statistika adalah 60, untuk menguji asumsi tersebut diambil sampel sebanyak 10 mahasiswa. Ujilah apakah asumsi tersebut apakah terdapat perbedaan yang signifikan antara rata –rata sampel dengan rata-rata populasi. Gunakan taraf signifikan alpha 0.05

X	80	90	60	60	90	60	80	50	70	60
---	----	----	----	----	----	----	----	----	----	----

## B. UJI t DUA SAMPEL

### 1. Uji t Sampel Dependen (Berpasangan)

Uji t berpasangan (*paired t-test*) biasanya menguji perbedaan antara dua pengamatan. Uji t berpasangan biasa dilakukan pada subjek yang diuji pada situasi sebelum dan sesudah proses, atau subjek yang berpasangan. Uji t digunakan untuk membandingkan mean dari suatu sampel yang berpasangan (*paired*). Sampel berpasangan adalah sebuah kelompok sampel dengan subyek yang sama namun mengalami dua perlakuan atau pengukuran yang berbeda.

Kriteria data untuk uji t sampel berpasangan :

- ) Data untuk tiap pasang yang diuji dalam skala interval atau rasio.
- ) Data berdistribusi normal.
- ) Nilai variannya dapat sama ataupun tidak.

Uji-t berpasangan adalah salah satu metode pengujian hipotesis dimana data yang digunakan tidak bebas (berpasangan). Uji-t ini membandingkan satu kumpulan pengukuran yang kedua dari contoh yang

sama. Uji ini sering digunakan untuk membandingkan skor sebelum dan sesudah percobaan untuk menentukan apakah perubahan nyata telah terjadi. Ciri-ciri yang paling sering ditemui pada kasus yang berpasangan adalah satu individu (objek penelitian) dikenai 2 buah perlakuan yang berbeda. Walaupun menggunakan individu yang sama, peneliti tetap memperoleh 2 macam data sampel, yaitu data dari perlakuan pertama (sebelum) dan data dari perlakuan kedua (sesudah). Perlakuan pertama mungkin saja berupa kontrol, yaitu tidak memberikan perlakuan sama sekali terhadap objek penelitian. Misal pada penelitian mengenai efektivitas suatu metode belajar, perlakuan pertama, peneliti menerapkan kontrol, sedangkan pada perlakuan kedua, barulah objek penelitian dikenai suatu tindakan tertentu, misal memberikan pelatihan. Dengan demikian, performance pelatihan dapat diketahui dengan cara membandingkan kondisi objek penelitian sebelum dan sesudah diberikan pelatihan.

#### Uji-t Berpasangan digunakan :

Dalam melakukan pemilihan uji, seorang peneliti harus memperhatikan beberapa aspek yang menjadi syarat sebuah uji itu digunakan. Peneliti tidak boleh sembarangan dalam memilih uji, sehingga sesuai dengan tujuan penelitian yang diinginkan. Ada beberapa syarat yang harus dipenuhi untuk menggunakan Uji-t Berpasangan. Dalam hal ini untuk Uji Komparasi antar dua nilai pengamatan berpasangan, (*paired*) misalnya sebelum dan sesudah (Pretest & posttest) di gunakan pada :

1. Satu sampel (setiap elemen ada 2 pengamatan)
2. Data kuantitatif (interval – rasio)
3. Berasal dari populasi yang berdistribusi normal (di populasi terdapat distribusi deference =  $d$  yang berdistribusi normal dengan mean  $md = 0$  dan variance  $sd^2 = 1$ ).

Setelah data yang dimiliki memenuhi syarat diatas, maka pemilihan uji statistik harus memperhatikan pertanyaan dari penelitian. Setelah melihat pertanyaan penelitian seorang peneliti kemudian melakukan pemilihan uji yang tepat untuk

menganalisis data yang dimiliki untuk menjawab pertanyaan penelitian yang disusun. Beberapa contoh pertanyaan penelitian yang menggunakan analisis / Uji t Berpasangan :

- ) Apakah ada perbedaan hasil belajar sebelum dan setelah melakukan pelatihan ?
- ) Apakah ada perubahan kualitas roti setelah terigu disubstitusi 30% dengan tepung ubi jalar?
- ) Apakah ada perbedaan kadar kolesterol dalam darah (mg %) yang di periksa oleh dua alat yang berbeda ?

**Rumus**

**D**

**t = -----**

**(SD/ N)**

No elemen	Observasi 1 Setelah (A <sub>i</sub> )	Observasi 2 Sebelum (B <sub>i</sub> )	Di (A <sub>i</sub> - B <sub>i</sub> )	Di <sup>2</sup> (A <sub>i</sub> - B <sub>i</sub> ) <sup>2</sup>
1	A <sub>1</sub>	B <sub>1</sub>	(A <sub>1</sub> - B <sub>1</sub> )	(A <sub>1</sub> - B <sub>1</sub> ) <sup>2</sup>
2	A <sub>2</sub>	B <sub>2</sub>	(A <sub>2</sub> - B <sub>2</sub> )	(A <sub>2</sub> - B <sub>2</sub> ) <sup>2</sup>
.	.	.	.	.
.	.	.	.	.
N	A <sub>n</sub>	B <sub>n</sub>	(A <sub>n</sub> - B <sub>n</sub> )	(A <sub>n</sub> - B <sub>n</sub> ) <sup>2</sup>
			Di	Di <sup>2</sup>

Rata-rata D =  $\frac{\sum Di}{n}$

Simpangan baku SD =  $\sqrt{\frac{\sum Di^2 - (\sum Di)^2/n}{n-1}}$

**Uji Hipotesis :**

Statistik hitung ( t hitung)

$$t = \frac{D}{SD / \sqrt{n}}$$

Titik kritis statistik t (t tabel) dilihat di tabel distribusi sampling t, dengan derajat bebas  $df = n-1$

Hipotesis  $H_0$  di tolak jika  $t_{hitung} > t_{tabel}$

Hipotesis  $H_0$  di terima jika  $t_{hitung} < t_{tabel}$

### Contoh kasus:

Berikut ini adalah contoh penelitian yang menggunakan uji-t berpasangan yaitu peneliti bermaksud meneliti perbedaan hasil belajar sebelum melakukan latihan intensif selama 7 hari berturut-turut dengan setelah melakukan latihan. Dengan sampel acak (*random*) terdiri dari 5 mahasiswa ditest sebelum dan sesudah melakukan latihan intensif.

Mahasiswa					
	1	2	3	4	5
Pretest	60	65	60	65	65
Posttest	70	75	75	65	60

Setelah dilakukan uji normalitas, ternyata populasi asal sampel mempunyai distribusi **normal**.

### Pertanyaan penelitian

Dalam penelitian ini pertanyaan penelitiannya adalah “Apakah terdapat perbedaan hasil belajar antara sebelum dan sesudah latihan intensif (dengan  $\alpha = 0,05$ )?”. Dengan demikian penelitian di atas menggunakan uji-t berpasangan karena setelah dilihat syarat-syaratnya telah dipenuhi untuk dilakukan uji-t berpasangan. Kemudian untuk melakukan uji tersebut ada beberapa langkah-langkah yang harus dilakukan. Berikut langkah-langkah untuk melakukan pengujian hipotesis :

#### 1. Rumuskan hipotesis:

$H_0$ : Tidak terdapat perbedaan hasil belajar antara sebelum dan sesudah latihan intensif

$H_1$  : Terdapat perbedaan hasil belajar antara sebelum dan sesudah latihan intensif

**2. Tabel Data dan perhitungan uji t**

No.	Hasil belajar		Di	Di <sup>2</sup>
	Pretest	Posttest		
1	60	70	10	100
2	65	75	10	100
3	60	75	15	225
4	65	65	0	0
5	65	60	-5	25
Jumlah			30	450

Rata-rata  $D = 30/5 = 6$

Simpangan baku  $SD = \frac{450-(30)^2 / 5}{5 - 1} = 8,22$

**t hitung** :  $t = \frac{D}{SD / n}$

$t = \frac{6}{8,22 / 5} = 1,63$

**3. Lihat tabel t untuk menetapkan nilai kritis t (t tabel):**

Langkah selanjutnya adalah dengan melihat t tabel untuk kemudian di bandingkan dengan hasil perhitungan yang dilakukan (t hitung).

Pada tingkat kemaknaan ( ) = 5% (0,05), dengan **df=n-1** = 5-1 =4

**Tabel (t-tabel)**

Df	Level of significance for one-tailed test					
	.10	.05	.025	.01	.005	.0005

	Level of significance for two-tailed test					
	.20	.10	.05	.02	.01	.005
1	-.-	-.-	12.706	-.-	-.-	-.-
2	-.—	-.—	4.303	-.—	-.—	-.—
3	-.—	-.—	3.182	-.—	-.—	-.—
4	-.—	-.—	<b>2.776</b>	-.—	-.—	-.—
5	-.—	-.—	2.571	-.—	-.—	-.—
6	-.—	-.-	2.447	-.-	-.-	-.-
7	-.-	-.-	2.365	-.-	-.-	-.-
8	-.—	-.—	2.306	-.—	-.—	-.—
9	-.—	-.—	2.262	-.—	-.—	-.—
10	-.—	-.—	2.228	-.—	-.—	-.—
11	-.—	-.—	2.201	-.—	-.—	-.—
12	-.-	-.-	2.179	-.-	-.-	-.-

Setelah melihat tabel t ditemukan t tabelnya adalah 2,776 Karena  $t_{hitung} (1,63) < t_{tabel} (2,776)$ , maka **hipotesis  $H_0$  diterima**. Dengan demikian bisa disimpulkan bahwa “Tidak terdapat perbedaan frekuensi nadi antara sebelum dan sesudah latihan interval”

## 2. Uji Beda Dua Rata-Rata Independen

Uji ini untuk mengetahui perbedaan rata-rata dua populasi atau kelompok data yang independen. Contoh kasus suatu penelitian ingin mengetahui hubungan status merokok ibu hamil dengan berat badan bayi yang dilahirkan. Respondan terbagi dalam dua kelompok, yaitu mereka yang merokok dan yang tidak merokok.

Syarat atau asumsi yang harus dipenuhi yaitu

- Data berdistribusi normal
- Kedua kelompok data independen

- o Variabel berbentuk numerik dan kategori (variabel kategori hanya dengan 2 kelompok)

Prinsip pengujian dua rata-rata adalah melihat perbedaan variasi kedua kelompok data sehingga diperlukan informasi apakah varian ke dua kelompok yang diuji sama atau tidak. Karena bentuk varian ke dua kelompok data berpengaruh pada nilai standar error

Langkah-langkah:

- Uji kesamaan varian. Jika variannya sama, maka dilakukan Uji-t independen dengan asumsi varian sama
- Uji-t independen. Jika variannya tidak sama, dilakukan Uji-t independen dengan asumsi varian tidak sama

Bentuk uji untuk varian yang sama

$$t = \frac{X_1 - X_2}{SD \sqrt{(1/n_1) + (1/n_2)}}$$

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

- $df = n_1 + n_2 - 2$

$n_1$  atau  $n_2$  = jumlah sampel kelompok 1 atau 2

$s_1$  atau  $s_2$  = standar deviasi sampel kelompok 1 dan 2

Bentuk uji untuk varian yang berbeda

$$t = \frac{X_1 - X_2}{\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}}$$

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

Uji homogenitas varian

Uji homogenitas digunakan untuk mengetahui varian antara kelompok data ke-1 apakah sama dengan kelompok data ke -2. Hal ini dilakukan dengan menggunakan uji F

$$F = \frac{S_1^2}{S_2^2}$$

Dimana  $df_1 = n_1 - 1$

$df_2 = n_2 - 1$



# MATERI KULIAH: STATISTIKA

**Bahan Kajian** : Uji anova  
**Program Studi** : Pendidikan Teknik Informatika  
**Minggu ke-** : 15-16

## A. Pembelajaran :

### 1. Learning Outcomes

Memiliki kemampuan tentang uji Anova

### 2. Materi

Anova one way

### 3. Uraian Materi

#### 1. PENDAHULUAN

Setiap perusahaan/lembaga perlu melakukan pengujian terhadap kumpulan hasil pengamatan mengenai suatu hal, skor hasil belajar siswa, berat bayi yang baru lahir, misalnya hasil penjualan produk, hasil produksi, gaji pekerja di suatu perusahaan nilainya bervariasi antara satu dengan yang lainnya. Hal ini berhubungan dengan varian dan rata-rata yang banyak digunakan untuk membuat kesimpulan melalui penaksiran dan pengujian hipotesis mengenai parameter, maka dari itu dilakukan analisis varian yang ada dalam cabang ilmu statistika industri yaitu ANOVA.

Varians untuk sekumpulan data melukiskan derajat perbedaan atau variasi nilai data individu yang ada dalam kelompok data tersebut. Secara umum varians dapat digolongkan ke dalam varians sistematis dan varians galat. Varians sistematis adalah pengukuran karena adanya pengaruh yang menyebabkan skor atau nilai data lebih condong ke satu arah tertentu dibandingkan ke arah lain.

Salah satu jenis varians sistematis dalam kumpulan data hasil penelitian adalah varians antar kelompok atau disebut juga varians eksperimental. Varians ini menggambarkan adanya perbedaan antara kelompok-kelompok hasil pengukuran. Dengan demikian varians ini terjadi karena adanya perbedaan antara kelompok-kelompok individu.

Jika uji kesamaan dua rata-rata atau uji t digunakan untuk mencari perbedaan atau persamaan dua rata-rata, maka uji beberapa rata-rata digunakan untuk mencari perbedaan atau persamaan beberapa rata-rata. Uji ini disebut dengan nama **analysis of variance (anova atau anava)**.

## 2. PENGERTIAN ANOVA

Anava atau Anova adalah sinonim dari analisis varians terjemahan dari *analysis of variance*, sehingga banyak orang menyebutnya dengan anova. Anova merupakan bagian dari metoda analisis statistika yang tergolong analisis komparatif lebih dari dua rata-rata

Analisis Varians (ANAVA) adalah teknik analisis statistik yang dikembangkan dan diperkenalkan pertama kali oleh Sir R. A Fisher. ANAVA dapat juga dipahami sebagai perluasan dari uji-t sehingga penggunaannya tidak terbatas pada pengujian perbedaan dua buah rata-rata populasi, namun dapat juga untuk menguji perbedaan tiga buah rata-rata populasi atau lebih sekaligus.

Jika menguji hipotesis nol bahwa rata-rata dua buah kelompok tidak berbeda, teknik ANAVA dan uji-t (uji dua pihak) akan menghasilkan kesimpulan yang sama; keduanya akan menolak atau menerima hipotesis nol. Dalam hal ini, statistik F pada derajat kebebasan 1 dan n-k akan sama dengan kuadrat dari statistik t.

ANAVA digunakan untuk menguji perbedaan antara sejumlah rata-rata populasi dengan cara membandingkan variansinya. Pembilang pada rumus variansi tidak lain adalah jumlah kuadrat skor simpangan dari rata-ratanya, yang secara sederhana dapat ditulis sebagai  $\sum(X_i - \mu)^2$ . Istilah jumlah kuadrat skor simpangan sering disebut jumlah kuadrat (sum of squares). Jika jumlah kuadrat tersebut dibagi dengan n atau n-1 maka akan diperoleh rata-rata kuadrat yang tidak lain dari variansi suatu distribusi. Rumus untuk menentukan varians sampel yaitu,

$$S^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n - 1}$$

Seandainya kita mempunyai suatu populasi yang memiliki variansi  $\sigma^2$  dan rata-rata  $\mu$ . Dari populasi tersebut misalkan diambil tiga buah sampel secara independent, masing-masing dengan  $n_1$ ,  $n_2$ , dan  $n_3$ . Dari setiap sampel tersebut dapat ditentukan rata-rata dan variansinya, sehingga akan diperoleh tiga buah rata-rata dan variansi sampel yang masing-masing merupakan statistik (penaksir) yang tidak bias bagi parameterinya. Dikatakan demikian karena, dalam jumlah sampel yang tak hingga, rata-rata dari rata-rata sampel akan sama dengan rata-rata populasi ( $\mu$ ) dan rata-rata dari variansi sampel juga akan sama dengan variansi populasi ( $\sigma^2$ ).

Ada dua hal yang perlu diperhatikan, yaitu:

1. Apabila memiliki 3 buah variansi sampel ( $S_i^2$ ) yang masing-masing merupakan penaksir yang tidak bias bagi variansi populasinya. Jika  $n_1=n_2=n_3=\dots=n_k$ , maka seluruh variansi sampel tersebut dapat dijumlahkan dan kemudian dibagi dengan banyaknya sampel ( $k$ ) sehingga akan diperoleh rata-rata variansi sampel yang dalam jangka panjang akan sama dengan variansi populasi. Dalam bahasa ANAVA, rata-rata variansi sampel ini dikenal dengan rata-rata jumlah kuadrat dalam kelompok (RJKD) atau mean of squares within groups ( $MS_w$ ).
2. Apabila memiliki 3 buah rata-rata sampel yang dapat digunakan untuk menentukan rata-rata dari rata-rata sampel. Simpangan baku distribusi rata-rata sampel ( $S_{\bar{Y}}$ ) atau galat baku rata-rata adalah simpangan baku distribusi skor dibagi dengan akar pangkat dua dari besarnya sampel.

$$S_{\bar{Y}} = \frac{S_Y}{\sqrt{n}}$$

Sejalan dengan itu, variansi distribusi rata-rata sampel  $S_{\bar{Y}}^2$  dapat ditulis sebagai berikut:

$$S_{\bar{Y}}^2 = \frac{S^2}{n}$$

Dengan demikian,  $S^2$  sebagai penaksir yang tidak bias bagi variansi populasi akan ekuivalen dengan variansi distribusi rata-rata dikalikan dengan besarnya sampel ( $n$ ) yang secara aljabar dapat ditulis sebagai berikut.

$$nS_{\bar{y}}^2 = S^2$$

Dalam konteks ANAVA,  $nS_{\bar{y}}^2$  dikenal dengan sebutan rata-rata jumlah kuadrat antar kelompok (RJKA) atau mean of squares between groups ( $MS_B$ ). Jika seluruh sampel diambil secara acak dari populasi yang sama, maka  $MS_B=MS_W$  atau RJKA = RJKD,

Sehingga,

$$F = MS_B / MS_W = RJKA / RJKD = \frac{\sigma^2}{\sigma^2} = 1$$

ANAVA digunakan untuk menguji hipotesis nol tentang perbedaan dua buah rata-rata atau lebih. Secara formal, hipotesis tersebut dapat ditulis sebagai berikut.

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

$$H_1: P \quad t_i \quad s_i \quad h \quad s_i \quad t_i \quad s_i \quad d \quad (=) t_i \quad b$$

Hipotesis nol di atas mengatakan bahwa rata-rata populasi pertama sama dengan rata-rata populasi ke dua dan seterusnya yang berarti bahwa seluruh sampel diambil dari populasi yang sama. Jika demikian maka, rata-ratanya akan mirip satu sama lain. Dalam menguji hipotesis nol tersebut, ANAVA melakukan perbandingan antara variansi antar kelompok ( $MS_B$ ) dengan variansi dalam kelompok ( $MS_W$ ). Jika ternyata kedua variansi itu sama ( $F=1$ ) maka berarti seluruh sampel yang dianalisis berasal dari populasi yang sama, dan kita tidak memiliki dasar untuk menolak hipotesis nol. Namun, jika ada salah satu nilai rata-rata yang jauh berbeda dengan nilai rata-rata lainnya maka berarti sampel tersebut berasal dari populasi yang berbeda.

Seluruh subjek yang berada dalam satu kelompok memiliki karakteristik yang sama pada peubah bebas yang tengah dikaji. Dalam bahasa eksperimen, mereka seluruhnya menerima perlakuan yang sama, sehingga keragaman mereka pada peubah terikat dipandanga sebagai keragaman galat dan tidak berkaitan dengan perbedaan jenis perlakuan atau peubah bebas.

Perbedaan rata-rata antar kelompok terdiri atas dua unsur yaitu keragaman galat dan keragaman yang berkaitan perbedaan pada peubah bebas. Oleh karena

keragaman di dalam kelompok ( $MS_W$ ) merupakan penaksir yang tidak bias atas variansi populasi dan keragaman antara kelompok ( $MS_B$ ) terdiri atas  $MS_W$  dan keragaman yang berkaitan dengan perlakuan, maka hubungan antara keduanya dapat dituliskan sebagai berikut:

$$MS_W = \sigma^2$$

$$MS_B = \sigma^2 + d\tau \quad p$$

Dengan demikian, F dapat juga dituliskan:

$$F = MS_B / MS_W$$

$$F = (\sigma^2 + d\tau \quad p \quad ) / \sigma^2$$

Jika dampak perlakuan sama dengan nol, maka

$$F = \frac{\sigma^2}{\sigma^2} = 1$$

Persoalan sekarang adalah bagaimana membedakan pengaruh yang sistematis dari pengaruh yang tidak sistematis (acak). ANAVA dan statistika inferensial pada umumnya mendekati persoalan ini dengan menggunakan teori peluang. Statistika inferensial bertugas untuk menjawab suatu pertanyaan yang dapat dirumuskan sebagai berikut: "jika hipotesis nol ternyata benar berapakah peluang memperoleh harga statistik tertentu?"

Misalkan dalam ANAVA, kita memperoleh  $F = 3,96$ . Pertanyaan yang harus dijawab adalah "berapa besar peluang memperoleh  $F = 3,96$  jika ternyata hipotesis nol itu benar?" Paket analisis statistik pada komputer umumnya memberikan jawaban terhadap pertanyaan tersebut secara langsung dalam bentuk  $p = 0,25, 0,01, 0,001$  dan sebagainya, namun jika dilakukan secara manual maka harga  $F_{hitung}$  harus dibandingkan dengan nilai kritis yang sudah disediakan dalam bentuk  $F_{tabel}$  pada derajat kebebasan dan tingkat keyakinan. Nilai  $p$  yang lebih kecil dari nilai yang ditentukan menunjukkan penolakan terhadap  $H_0$ . Kesimpulan yang sama diperoleh jika ternyata  $F_{hitung} > F_{tabel}$ . Menolak hipotesis nol berarti menyimpulkan bahwa perbedaan antara  $MS_B$  dengan  $MS_W$  berkaitan dengan pengaruh yang sistematis dari faktor atau peubah bebas yang diteliti.

### 3. ANOVA SATU ARAH

Dinamakan analisis varians satu arah atau satu jalur, karena analisisnya menggunakan varians dan data hasil pengamatan merupakan pengaruh satu faktor. Dari tiap populasi secara independen diambil sebuah sampel acak, berukuran  $n_1$  dari populasi kesatu,  $n_2$  dari populasi kedua dan seterusnya berukuran  $n_k$  dari populasi ke k. Data sampel akan dinyatakan dengan  $Y_{ij}$  yang berarti data ke-j dalam sampel yang diambil dari populasi ke-i.

ANAVA satu arah atau jalur yaitu analisis yang melibatkan hanya satu peubah bebas. Secara rinci, ANAVA satu jalur digunakan dalam suatu penelitian yang memiliki ciri-ciri berikut yaitu melibatkan hanya satu peubah bebas dengan dua kategori atau lebih yang dipilih dan ditentukan oleh peneliti secara tidak acak. Kategori yang dipilih disebut tidak acak karena peneliti tidak bermaksud menggeneralisasikan hasilnya ke kategori lain di luar yang diteliti pada peubah itu. Sebagai contoh, peubah jenis kelamin hanya terdiri atas dua ketegori (pria-wanita), atau peneliti hendak membandingkan keberhasilan antara Metode A, B, dan C dalam meningkatkan semangat belajar tanpa bermaksud mengenerelisasikan ke metode lain di luar ketiga metode tersebut.

1. Perbedaan antara kategori atau tingkatan pada peubah bebas dapat bersifat kualitatif atau kuantitatif.
2. Setiap subjek merupakan anggota dari hanya satu kelompok pada peubah bebas, dan dipilih secara acak dari populasi tertentu

Sedangkan keterangan istilah yang sering digunakan dalam analisa varian yaitu faktor, level faktor dan variabel dependen. Faktor adalah variabel independen yang dipelajari dalam suatu penelitian. Level faktor, adalah bentuk/kondisi khusus dari suatu faktor. Variabel dependen, adalah variabel yang dipengaruhi oleh faktor.

Tujuan dari uji anova satu arah adalah untuk membandingkan lebih dari dua rata-rata. Sedangkan gunanya untuk menguji kemampuan generalisasi. Maksudnya dari signifikansi hasil penelitian. Jika terbukti berbeda berarti kedua sampel tersebut dapat digeneralisasikan (data sampel dianggap dapat mewakili populasi). Anova satu jalur dapat melihat perbandingan lebih dari dua kelompok data.

Anova pengembangan atau penjabaran lebih lanjut dari uji-t (  $t_{hit}$  ) .Uji-t atau uji-z hanya dapat melihat perbandingan dua kelompok data saja. Sedangkan anova satu jalur lebih dari dua kelompok data. Contoh: Perbedaan prestasi belajar statistika antara mahasiswa tugas belajar ( $X_1$ ), izin belajar ( $X_2$ ) dan umum ( $X_3$ ).

Anova lebih dikenal dengan uji-F (*Fisher Test*), sedangkan arti variasi atau varian itu asalnya dari pengertian konsep “*Mean Square*” atau kuadrat rerata (KR).

Rumusnya :

$$K = \frac{J}{d}$$

Dimana:  $J$  = jumlah kuadrat (*some of square*)

$d$  = derajat bebas (*degree of freedom*)

Menghitung nilai Anova atau F (  $F_{hit}$  ) dengan rumus :

$$F_{hit} = \frac{V_A}{V_D} = \frac{K_A}{K_D} = \frac{J_A : d_A}{J_D : d_D} = \frac{v}{v} \frac{a}{a} \frac{g}{g}$$

Varian dalam group dapat juga disebut Varian Kesalahan (Varian Galat). Dapat dirumuskan :

$$JK_A = \frac{(\sum X_A)^2}{n_A} - \frac{(\sum X_T)^2}{N} \text{ untuk } d_A = A - 1$$

$$JK_D = (\sum X_T)^2 - \sum \frac{(\sum X_A)^2}{n_A} \text{ untuk } d_D = N - A$$

Dimana

$$\frac{(\sum X_T)^2}{N} = \text{sebagai faktor koreksi}$$

N = Jumlah keseluruhan sampel (jumlah kasus dalam penelitian).

A = Jumlah keseluruhan group sampel.

### Langkah-langkah Anova Satu Arah

Prosedur Uji Anova Satu Arah

- 1) Sebelum anova dihitung, asumsikan bahwa data dipilih secara random, berdistribusi normal, dan variannya homogen.

Jika asumsi di atas tidak terpenuhi, maka ada prosedur khusus yang harus dilakukan. Misalnya, data yang dianalisis tidak mengikuti distribusi normal. Penanganan yang dapat dilakukan untuk masalah tersebut salah satunya adalah dengan mentransformasi data. Selain dengan transformasi data, cara lain untuk menganalisis data tersebut adalah dengan melakukan uji non parametrik dimana uji tersebut tidak membutuhkan asumsi-asumsi seperti pada uji parametrik. Uji non parametrik yang dapat dilakukan misalnya : uji Mann-Whitney, Wilcoxon signed-rank test, Sign test, Kruskal Wallis, dan lain-lain.

- 2) Buatlah hipotesis ( $H_a$  dan  $H_0$ ) dalam bentuk kalimat.
- 3) Buatlah hipotesis ( $H_a$  dan  $H_0$ ) dalam bentuk statistik.
- 4) Buatlah daftar statistik induk.
- 5) Hitunglah jumlah kuadrat antar group ( $J_A$ ) dengan rumus :

$$J_A = \frac{(\sum X_A)^2}{n_A} - \frac{(\sum X_t)^2}{N} = \left( \frac{(\sum X_{A1})^2}{n_{A1}} + \frac{(\sum X_{A2})^2}{n_{A2}} + \frac{(\sum X_{A3})^2}{n_{A3}} \right) - \frac{(\sum X_t)^2}{N}$$

- 6) Hitunglah derajat bebas antar group dengan rumus :  $d_A = A - 1$
- 7) Hitunglah kudrat rerata antar group ( $K_A$ ) dengan rumus :  $K_A = \frac{J_A}{d_A}$
- 8) Hitunglah jumlah kuadrat dalam antar group ( $J_D$ ) dengan rumus :

$$J_D = (\sum X_t)^2 - \sum \frac{(\sum X_A)^2}{n_A}$$

$$= \sum X_{A1}^2 + \sum X_{A2}^2 + \sum X_{A3}^2 - \left( \frac{(\sum X_{A1})^2}{n_{A1}} + \frac{(\sum X_{A2})^2}{n_{A2}} + \frac{(\sum X_{A3})^2}{n_{A3}} \right)$$

- 9) Hitunglah derajat bebas dalam group dengan rumus :  $d_D = N - A$
- 10) Hitunglah kudrat rerata dalam antar group ( $K_D$ ) dengan rumus :  $K_D = \frac{J_D}{d_D}$
- 11) Carilah  $F_{hit}$  dengan rumus :  $F_{hit} = \frac{K_A}{K_D}$
- 12) Tentukan taraf signifikansinya, misalnya  $\alpha = 0,05$  atau  $\alpha = 0,01$
- 13) Cari  $F_{t_i}$  dengan rumus :  $F_{t_i} = F_{(1-\alpha)(d_A, d_D)}$
- 14) Buat Tabel Ringkasan Anova



**Tabel. Ringkasan Anova Satu Arah**

Sumber Varian (SV)	Jumlah Kuadrat (JK)	Derajat bebas (db)	Kuadrat Rerata (KR)	$F_{hit}$	Taraf Signifikan ( $\rho$ )
Antar group (A)	$\frac{(\sum X_A)^2}{n_A} - \frac{(\sum X_t)^2}{N}$	$A - 1$	$\frac{J_A}{d_A}$	$\frac{K_A}{K_D}$	$\alpha$
Dalam group (D)	$(\sum X_t)^2 - \sum \frac{(\sum X_A)^2}{n_A}$	$N - A$	$\frac{J_D}{d_D}$	-	-
Total	$(\sum X_t)^2 - \frac{(\sum X_t)^2}{N}$	$N - 1$	-	-	-

15) Tentukan kriteria pengujian : jika  $F_{hit} > F_t$  , maka tolak  $H_0$  berarti signifikan dan konsultasikan antara  $F_{hit}$  dengan  $F_t$  kemudian bandingkan

16) Buat kesimpulan.

**Contoh 1 Soal dan Pembahasan**

Seorang ingin mengetahui perbedaan prestasi belajar untuk mata kuliah dasar-dasar statistika antara mahasiswa tugas belajar, izin belajarn dan umum.

Data diambil dari nilai UTS sebagai berikut :

Tugas belajar ( $A_1$ ) = 6 8 5 7 7 6 6 8 7 6 7 = 11 orang

Izin belajar ( $A_2$ ) = 5 6 6 7 5 5 5 6 5 6 8 7 = 12 orang

Umum ( $A_3$ ) = 6 9 8 7 8 9 6 6 9 8 6 8 = 12 orang

Buktikan apakah ada perbedaan atau tidak?

**LANGKAH-LANGKAH MENJAWAB :**

1. Diasumsikan bahwa data dipilih secara random, berdistribusi normal, dan variannya homogen.
2. Hipotesis ( $H_a$  dan  $H_0$ ) dalam bentuk kalimat.

$H_a$  = Terdapat perbedaan yang signifikan antara mahasiswa tugas belajar, izin belajar dan umum.

$H_0$  = Tidak ada perbedaan yang signifikan antara mahasiswa tugas belajar, izin belajar dan umum.

3. Hipotesis ( $H_a$  dan  $H_0$ ) dalam bentuk statistic

$$H_a : A_1 = A_2 = A_3 \quad H_0 : A_1 \neq A_2 \neq A_3$$

4. Daftar statistik induk

NILAI UTS			
NO	$A_1$	$A_2$	$A_3$
1	6	5	6
2	8	6	9
3	5	6	8
4	7	7	7
5	7	5	8
6	6	5	9
7	6	5	6
8	8	6	6
9	7	5	9
10	6	6	8
11	7	8	6
12	-	7	8

STATISTIK				TOTAL(T)
$n$	11	12	12	N=35
$x$	73	71	90	234
$x^2$	943	431	692	1616
$\bar{X}$	6,64	5,92	7,5	6,69
$(\sum x)^2/n$	484,45	420,08	675	1564,46
Varians ( $S^2$ )	0,85	0,99	1,55	1,33

5. Menghitung jumlah kuadrat antar group ( $J_{iA}$ ) dengan rumus :

$$J_{iA} = \frac{(\sum X_A)^2}{n_A} - \frac{(\sum X_T)^2}{N}$$

$$= \left( \frac{(7)^2}{1} + \frac{(7)^2}{1} + \frac{(9)^2}{1} \right) - \frac{(2)^2}{3} = 1579,53 - 1564,46 = 15,07$$

6. Hitunglah derajat bebas antar group dengan rumus :

$$d_A = A - 1 = 3 - 1 = 2 \quad A = \text{jumlah group A}$$

7. Hitunglah kudrat rerata antar group ( $K_A$ ) dengan rumus :

$$K_A = \frac{J_A}{d_A} = \frac{1,0}{2} = 7,54$$

8. Hitunglah jumlah kuadrat dalam antar group ( $J_{iD}$ ) dengan rumus :

$$J_{iD} = (\sum X_T)^2 - \sum \frac{(\sum X_A)^2}{n_A} = (493 + 431 + 692) - \left( \frac{(7)^2}{1} + \frac{(7)^2}{1} + \frac{(9)^2}{1} \right)$$

$$= 1616 - 1579,53 = 36,47$$

9. Hitunglah derajat bebas dalam group dengan rumus :

$$d_D = N - A = 35 - 3 = 32$$

10. Hitunglah kudrat rerata dalam antar group ( $K_D$ ) dengan rumus :

$$K_D = \frac{J_D}{d_D} = \frac{3,4}{3} = 1,14$$

11. Carilah  $F_{hit}$  dengan rumus :

$$F_{hit} = \frac{K_A}{K_D} = \frac{7,54}{1,14} = 6,61$$

12. Tentukan taraf signifikansinya, misalnya  $\alpha = 0,05$

13. Cari  $F_{t_i}$  dengan rumus :

$$F_{t_i} = F_{(1-\alpha)(d_A, d_D)}$$

$$F_t = F_{(1-0,05)(2,3)}$$

$$F_t = F_{(0,95)(2,3)}$$

$$F_t = 3,30$$

14. Cara mencari : Nilai  $F_t = 3,30$  dan arti angka  $F_t = F_{(0,95)(2,3)}$

0,95 = Taraf kepercayaan 95% atau taraf signifikan 5%.

Angka 2 = pembilang atau hasil dari  $d_A$

Angka 32 = penyebut atau hasil dari  $d_D$

Apabila angka 2 dicari ke kanan dan angka 32 ke bawah maka akan bertemu dengan nilai  $F_t = 3,30$ . Untuk taraf signifikansi 5% dipilih pada bagian atas dan 1% dipilih pada bagian bawah.

Buat Tabel Ringkasan Anova

TABEL  
RINGKASSAN ANOVA SATU JALUR

Sumber Varian (SV)	Jumlah Kuadrat (JK)	Derajat Bebas (db)	Kuadrat Rerata (KR)	$F_{hit}$	Taraf Signifikan ( $\rho$ )
Antar group (A)	15,07	2	7,54	6,61	< 0,05 $F_t = 3,30$
Dalam group (D)	36,47	32	1,14	-	-
Total	51,54	54	-	-	-

15. Tentukan kriteria pengujian : jika  $F_{hit} > F_t$  , maka tolak  $H_0$  berarti signifikan. Setelah konsultasikan dengan tabel F kemudian bandingkan antara  $F_{hit}$  dengan  $F_t$  , ternyata :  $F_{hit} > F_t$  atau  $6,61 > 3,30$  maka tolak  $H_0$  berarti signifikan.
16. Kesimpulan  
 $H_0$  ditolak dan  $H_a$  diterima. Jadi, terdapat perbedaan yang signifikan antara mahasiswa tugas belajar, izin belajar dan umum.

### Contoh 2

Sebuah penelitian dilakukan untuk mengetahui apakah ada pengaruh perbedaan metode belajar pada tingkat prestasi siswa. Ada tiga metode belajar yang akan diuji. Diambil sampel masing-masing 5 guru untuk mengerjakan pekerjaannya, lalu dicatat waktu yang digunakan (menit) sebagai berikut:

Metode 1 (menit)	Metode 2 (menit)	Metode 3 (menit)
21	17	31
27	25	28
29	20	22
23	15	30
25	23	24

Ujilah dengan  $\alpha = 0,05$  apakah ada pengaruh perbedaan metode belajar pada waktu yang digunakan?

Penyelesaian :

Metode 1 (menit)	Metode 2 (menit)	Metode 3 (menit)
21	17	31
27	25	28
29	20	22
23	15	30

25	23	24
$T_1 = 125$	$T_2 = 100$	$T_3 = 135$

Dari tabel di atas bisa dihitung

Total keseluruhan nilai = 360

$$JKK = \frac{1^2}{5} + \frac{1^2}{5} + \frac{1^2}{5} - \frac{3^2}{1} = 130$$

$$JKT = 21^2 + 27^2 + \dots + 24^2 - \frac{3^2}{1} = 298$$

$$JKS = 298 - 130 = 168$$

Tabel ANOVA

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat	Varian (Ragam)	$F_{hitung}$	$F_{tabel}$
Antar Kolom	2	130	65	4,64	$F_{(2, 12)} = 3,89$
Sisaan	12	168	14		
	14	298			

Pengujian Hipotesis

$$H_0 = \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_a : \text{Tidak semuanya sama (s, a } \mu_i \neq \mu_j \text{ u } i \neq j)$$

$$\text{Statistik Uji} = F_{hitung} = 4,64$$

Karena  $F_{hitung} > F_{tabel}$  maka tolak  $H_0$

Kesimpulan: Ada pengaruh perbedaan metode kerja pada waktu yang digunakan.

### Contoh Soal 3

Dengan menggunakan Uji Anova Satu Jalur, selidikilah apakah ada perbedaan hasil belajar siswa dengan pembelajaran menggunakan metode ceramah (A), Tanya jawab (B), STAD (C), yang diberikan kepada 3 kelompok siswa dengan masing – masing kelompok terdiri dari 5 orang siswa.

Jawab :

$$H_0 : \mu_A = \mu_B = \mu_C$$

$$H_A : \mu_A \neq \mu_B \text{ atau } \mu_A \neq \mu_C \text{ atau } \mu_B \neq \mu_C$$

Metode Mengajar			A <sup>2</sup>	B <sup>2</sup>	C <sup>2</sup>
A	B	C			
5	6	7	25	36	49
6	7	8	36	49	64
6	7	8	36	49	64
6	7	8	36	49	64
7	8	9	49	64	81
A = 30	B = 35	C = 40	A <sup>2</sup> = 182	B <sup>2</sup> = 247	C <sup>2</sup> = 322

$$n = 5$$

$$\bar{A} = \frac{\sum A}{n_A} = \frac{30}{5} = 6$$

$$\bar{B} = \frac{\sum B}{n_B} = \frac{35}{5} = 7$$

$$\bar{C} = \frac{\sum C}{n_C} = \frac{40}{5} = 8$$

$$JK_t = \sum_{j=1}^k \sum_{i=1}^n X_{ij}^2 - \frac{J^2}{N}$$

$$JK_t = (25 + 36 + 36 + 36 + 49 + 36 + 49 + 49 + 49 + 64 + 49 + 64 + 64 + 64 + 81) -$$

$$\frac{(30 + 35 + 40)^2}{5}$$

$$JK_t = 751 - \frac{(105)^2}{5} = 751 - 735 = \mathbf{16}$$

$$JK_a = \sum_{j=1}^2 \frac{J_j^2}{n_j} - \frac{J^2}{N}$$

$$JK_a = \left( \frac{30^2}{5} + \frac{35^2}{5} + \frac{40^2}{5} \right) - 375$$

$$JK_a = 180 + 245 + 320 - 735 = \mathbf{10}$$

$$JK_i = JK_t - JK_a$$

$$JK_i = 16 - 10 = \mathbf{6}$$

Tabel Anova Satu Jalur

	JK	dk	RJK	F
Antar	10	2	5	$\frac{5}{0,5} = 10$
Inter	6	12	0,5	

F hitung = 10

Untuk  $\alpha = 0,05$  dan dk (2 , 12) diperoleh F tabel = 3,89 (lihat tabel)

Karena F hitung > F tabel , maka hipotesis ditolak, artinya ada perbedaan hasil belajar siswa dengan 3 metode mengajar yaitu metode ceramah, Tanya jawab dan STAD.



Tabel F untuk  $\alpha = 0,05$

df untuk penyebut (N2)	df untuk pembilang (N1)														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	161	199	216	225	230	234	237	239	241	242	243	244	245	245	246
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41	19.42	19.42	19.43
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.78	8.74	8.73	8.71	8.70
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.22	2.20	2.18
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.24	2.20	2.18	2.15	2.13
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.11
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.14	2.11	2.09
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13	2.10	2.08	2.06
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.08	2.05	2.03
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	2.01
31	4.16	3.30	2.91	2.68	2.52	2.41	2.32	2.25	2.20	2.15	2.11	2.08	2.05	2.03	2.00
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.10	2.07	2.04	2.01	1.99
33	4.14	3.28	2.89	2.66	2.50	2.39	2.30	2.23	2.18	2.13	2.09	2.06	2.03	2.00	1.98
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12	2.08	2.05	2.02	1.99	1.97
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	2.07	2.04	2.01	1.99	1.96
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11	2.07	2.03	2.00	1.98	1.95
37	4.11	3.25	2.86	2.63	2.47	2.36	2.27	2.20	2.14	2.10	2.06	2.02	2.00	1.97	1.95
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.99	1.96	1.94
39	4.09	3.24	2.85	2.61	2.46	2.34	2.26	2.19	2.13	2.08	2.04	2.01	1.98	1.95	1.93
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.97	1.95	1.92
41	4.08	3.23	2.83	2.60	2.44	2.33	2.24	2.17	2.12	2.07	2.03	2.00	1.97	1.94	1.92
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.03	1.99	1.96	1.94	1.91
43	4.07	3.21	2.82	2.59	2.43	2.32	2.23	2.16	2.11	2.06	2.02	1.99	1.96	1.93	1.91
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.95	1.92	1.90
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05	2.01	1.97	1.94	1.92	1.89

### **Daftar Pustaka**

- Sugiono.2010. Metode Penelitian Kuantitatif, kualitatif dan R & D.  
Suharsimi Arikunto.1987. Prosedur Penelitian Suatu Pendekatan Praktis.
- Nasir. 1983. Metode Penelitian. Ghalia Indonesia. Jakarta.
- Singgih Santoso. 2002. SPSS Statistik Multivariat. Elek Media Komputindo.  
Jakarta.
- Syamsul Ma'arif dan Hendri Tanjung. 2003. Teknik-Teknik kuantitatif untuk  
manajemen. Grasindo. Jakarta.