

BUKU AJAR

BIOSTATISTIK

DESKRIPTIF & INFERENSIAL

**Dr. MG. Catur Yuantari, S.KM, M.Kes
Sri Handayani, S.KM, M.Kes**

BUKU AJAR

Biostatistik Deskriptif & Inferensial

Dr. MG. Catur Yuantari, S.KM, M.Kes

Sri Handayani, S.KM, M.Kes



Diterbitkan oleh :

Badan Penerbit Universitas Dian Nuswantoro

ISBN : 979-26-0282-8

Cetakan ke-2 th 2017

KATA PENGANTAR

Dengan memanjatkan rasa syukur yang sedalam-dalamnya kehadirat Allah Yang Maha Kuasa, atas terselesaikannya buku ajar Biostatistik Deskriptif dan Inferensial. Buku ajar ini disusun dengan tujuan untuk memudahkan proses belajar mengajar di kelas, dimana mahasiswa mempunyai referensi dalam mempelajari mata kuliah Biostatistik Deskriptif dan Inferensial. Selain itu dengan menggunakan buku ajar, maka mahasiswa dapat menyerap materi dengan lebih baik, karena telah disesuaikan dengan satuan acara pengajaran.

Dalam penyusunan buku ini, penyusun menggunakan berbagai referensi, baik yang diambil secara langsung ataupun mengalami modifikasi dari penyusun. Oleh karena itu tetap diharapkan mahasiswa membaca referensi utama, yang ada pada setiap bab pembahasan.

Akhir kata, tidak ada gading yang tak retak, masih terdapat kekurangan untuk itu mohon kritik dan saran yang membangun. Mudah-mudahan buku ini dapat berguna dan bermanfaat bagi kita semua.

Semarang, Maret 2016

Penulis

DAFTAR ISI

Kata Pengantar	iii
Daftar Isi	iv
Pokok Bahasan 1 : PENGANTAR STATISTIKA	1
Pokok Bahasan 2 : PENGORGANISASIAN DATA	11
Pokok Bahasan 3 : DISTRIBUSI FREKUENSI	33
Pokok Bahasan 4 : UKURAN PEMUSATAN DATA	47
Pokok Bahasan 5 : UKURAN PENYEBARAN DATA	59
Pokok Bahasan 6 : PENDUGAAN PARAMETER.....	75
Pokok Bahasan 7 : HIPOTESIS	85
Pokok Bahasan 8 : VALIDITAS DAN RELIABILITAS	95
Pokok Bahasan 9 : UJI PARAMETRIK DAN NON PARAMETRIK	101
Pokok Bahasan 10 : UJI BEDA (T-test dan Anova)	107
Pokok Bahasan 11 : UJI KORELASI	129
Pokok Bahasan 12 : CHI SQUARE DAN FISHER EXACT	139
Pokok Bahasan 13 : REGRESI	143

1

Pokok Bahasan PENGANTAR STATISTIKA

1.1 PENDAHULUAN

1.1.1 Deskripsi Singkat

Pada pokok bahasan ini akan membahas tentang Pengantar Ilmu Statistika yang meliputi memahami sejarah, definisi, fungsi, manfaat hingga implementasi persoalan di masyarakat.

1.1.2 Relevansi

Materi dalam bab ini memberikan pemahaman bagi seorang sarjana kesehatan dalam menjelaskan pengertian statistik, statistika dan biostatistik, ruang lingkup biostatistik serta penerapan biostatistik dalam bidang kesehatan.

1.1.3 Kompetensi

a. Standar Kompetensi

Pokok bahasan ini memberikan kontribusi kompetensi kepada mahasiswa kesehatan masyarakat agar mampu menjelaskan pengertian statistik dan statistika; biostatistik, serta penerapan biostatistik pada bidang kesehatan.

b. Kompetensi Dasar

Setelah mengikuti materi Pengantar Statistika mahasiswa mampu Menjelaskan pengertian statistik &

statistika, menjelaskan pengertian biostatistik, menjelaskan sejarah biostatistik, menjelaskan ruang lingkup biostatistik, menjelaskan tujuan biostatistik, menjelaskan penerapan biostatistik pada bidang kesehatan, menjelaskan langkah-langkah prosedur Biostatistik.

1.2 PENYAJIAN

1.2.1 Uraian Isi

Dalam kehidupan sehari-hari kita telah terbiasa dengan kata “statistik”, walaupun dalam bentuk sederhana. Contohnya setiap bulan orang tua telah memberi uang bulanan untuk biaya hidup selama belajar sebesar Rp 1.500.000,- yang digunakan untuk biaya makan, membayar kost, membeli peralatan kuliah, dll. Namun banyak anggapan atau persepsi dari mahasiswa bahwa mata pelajaran statistika sangatlah sulit dan menakutkan. Hal ini berarti statistika dan matematika belum menjadi bagian dalam hidup. Mari kita belajar statistika dalam kehidupan sehari-hari untuk mempermudah pembelajaran.

A. PENGERTIAN STATISTIK DAN STATISTIKA

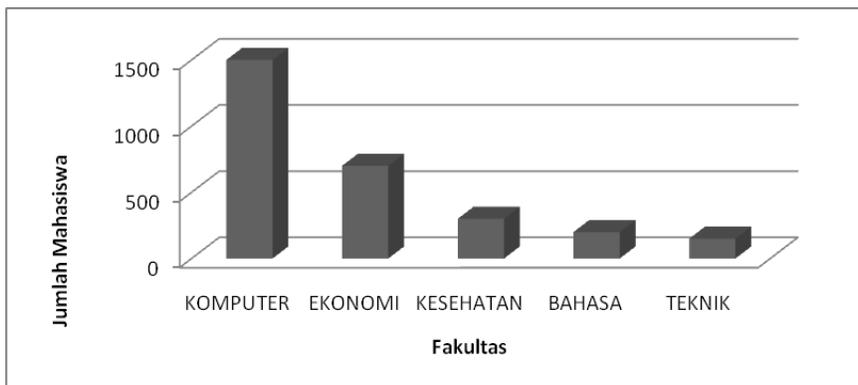
➤ STATISTIK

Statistik mempunyai beberapa pengertian, statistik dalam arti paling sederhana artinya data.

Contohnya:

1. Harga beras adalah Rp 10.000,- per kg.
2. Korban meningkat akibat tanah longsor mencapai 34 orang terdiri atas 12 orang laki-laki, 22 orang perempuan.

Pengertian statistik yang lebih luas adalah kumpulan data dalam bentuk angka maupun bukan angka yang disusun dalam tabel (daftar) dan atau diagram yang menggambarkan dengan suatu masalah tertentu. Kata statistik juga menyatakan ukuran atau karakteristik pada sampel seperti nilai rata-rata, standar deviasi, variasi dan koefisien korelasi. Pengertian statistik menurut UU RI no.7 tahun 1960 adalah keterangan berupa angka yang memberikan gambaran yang wajar dari seluruh ciri-ciri kegiatan atau keadaan masyarakat Indonesia. Contohnya : statistik penduduk adalah kumpulan angka-angka yang berkaitan dengan masalah penduduk; statistik kesehatan adalah kumpulan angka-angka yang berkaitan dengan masalah kesehatan.



Gambar : Diagram batang Jumlah Mahasiswa di Universitas Dian Nuswantoro

➤ STATISTIKA

Statistika adalah ilmu yang mempelajari cara untuk merencanakan, mengumpulkan, menganalisis, menginterpretasi dan mempresentasikan data. Statistika adalah pengetahuan yang berkaitan dengan metode, teknik atau cara untuk mengumpulkan data, mengolah data, menyajikan

data, menganalisis data dan menarik kesimpulan atau menginterpretasikan data.

Sebagian besar konsep dasar statistika mengasumsikan teori probabilitas. Beberapa istilah statistika meliputi populasi, sampel, unit, sampel dan probabilitas. Penerapan statistika dalam berbagai disiplin ilmu seperti ilmu alam dalam astronomi dan biologi, ilmu-ilmu sosial termasuk sosiologi dan psikologi, bidang kesehatan seperti kedokteran, kesehatan masyarakat maupun keperawatan.

“Statistik (tanpa huruf “a”) seringkali dikaburkan pengertiannya dengan “Statistika”. “Statistik” diartikan sebagai kumpulan angka hasil pengukuran atau perhitungan yang disebut dengan data. “Statistik” sering pula digunakan untuk menyatakan nilai hasil pengukuran atau perhitungan pada sebagian obyek pengamatan atau sampel sebagai pembeda dari “parameter”, yaitu suatu nilai yang diperoleh dari populasinya. Selanjutnya “statistika” diartikan pula sebagai metode atau alat bantu untuk mengembangkan ilmu pengetahuan melalui aktivitas berupa pengumpulan, pengolahan, penyajian dan analisis data yang dilanjutkan dengan penarikan kesimpulan atas ciri yang diamati dari sampel terhadap populasinya.

Dengan singkat dapat didefinisikan bahwa statistika adalah pengetahuan yang berkaitan dengan statistik. Oleh karena itu kita mengenal ilmu statistika bukan ilmu statistik.

Pengetahuan dan penerapan statistika banyak dipakai dalam metodologi penelitian karena penelitian merupakan serangkaian kegiatan yang meliputi mengumpulkan data, mengolah data, menyajikan data, menganalisis data menginterpretasikan dan menarik kesimpulan dari

sekumpulan data yang kemudian ditulis secara lengkap dan berurutan dalam bentuk laporan penelitian.

B. PERKEMBANGAN STATISTIKA

Metode Statistika telah dikenal sejak zaman Romawi kuno yang berasal dari bahasa Italia “Statista” yang berarti “negarawan”. Maksudnya penggunaannya terbatas untuk kepentingan Negara, misalnya berkaitan dengan penarikan pajak, wajib militer, dan lain-lain. Data yang diperlukan pada saat itu masih sangat terbatas, misalnya nama, umur, jenis kelamin, pekerjaan, dan jumlah keluarga. Metode statistik adalah serangkaian prosedur (Metode) untuk melakukan kegiatan statistik yang sistematis. Analisa statistik adalah kegiatan Analisa terhadap data kuantitatif untuk menarik kesimpulan.

Penggunaan statistika dalam bidang kesehatan diawali oleh John Graunt (1662) melalui pencatatan tentang kematian. Selanjutnya diikuti oleh sarjana-sarjana lain seperti William farr, Karl Pearson dan lain-lain.

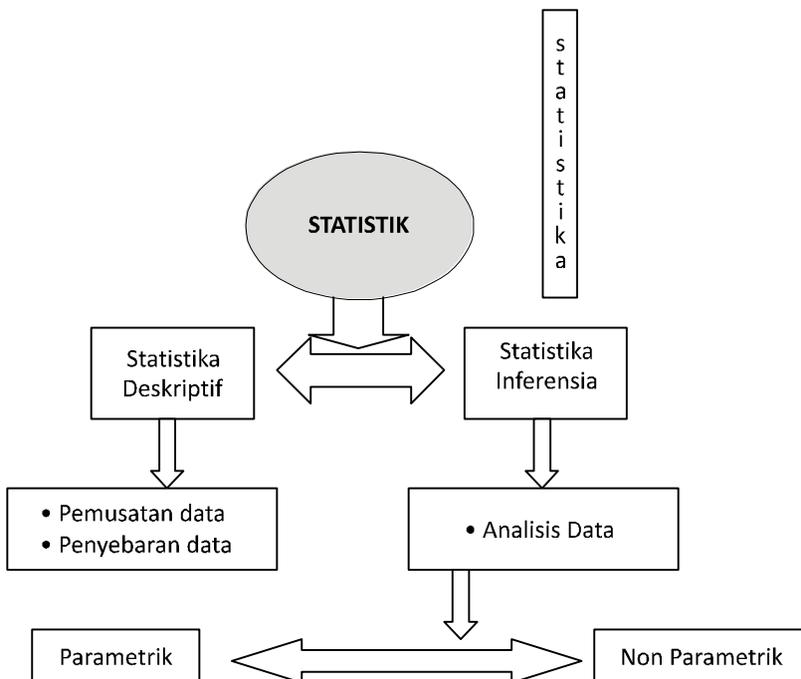
Statistika Kesehatan adalah ilmu terapan (applied science) metode statistik terhadap masalah kesehatan didalamnya termasuk Vital statistik yang membahas tentang komponen daur hidup (Life statistik), seperti statistik kematian, kelahiran, perkawinan dan lain-lain. *Health service* statistik pengolahan data statistik untuk mengukur output dari kegiatan pelayanan kesehatan.

C. PENGGOLONGAN STATISTIKA

Metode statistika telah digolongkan menjadi dua yaitu metode statistika deskriptif dan metode statistika inferensia.

Statistika deskriptif merupakan kegiatan mulai dari pengumpulan data sampai mendapatkan informasi dengan jalan menyajikan dan menganalisis data yang telah dikumpulkan. Informasi yang didapatkan dari statistika deskripsi seperti pemusatan data (mean, median, modus), penyebaran data (range, simpangan rata-rata, varians dan simpangan baku).

Statistika inferensia mempelajari cara menganalisis data serta mengambil kesimpulan (berkaitan dengan estimasi parameter dan pengujian hipotesis. Metode ini sering disebut statistika induktif karena kesimpulan yang ditarik didasarkan pada informasi dari sebagian data. Statistika inferensia dibagi dalam dua kelompok yaitu statistika parametrik dan non parametrik.



Statistik parametrik merupakan bagian dari statistika inferensia yang mempertimbangkan nilai dari satu atau lebih parameter populasi. Statistika parametrik biasanya dihubungkan dengan data kuantitatif serta mempunyai syarat berdistribusi normal. Sedangkan statistik nonparametrik merupakan bagian dari statistika inferensia yang tidak memperhatikan nilai dari satu atau lebih parametrik.

D. PENGERTIAN BIOSTATISTIK

Penggunaan metode statistik untuk memecahkan masalah kesehatan yang unsur utamanya adalah manusia dikenal dengan Biostatistik. Biostatistika merupakan ilmu terapan dari statistika dalam bidang biologi. Dalam kenyataannya Biostatistika juga banyak digunakan dalam bidang kesehatan dan kedokteran, karena keduanya memang terkait erat dengan bidang biologi. Sedangkan statistika sebagai cabang ilmu matematika banyak digunakan dalam pengambilan keputusan dan berkembang berdasarkan teori peluang (probabilitas)

RUANG LINGKUP BIOSTATISTIK :

- Medis
- Kependudukan
- Kesehatan lingkungan
- Kesehatan kerja
- Administrasi kesehatan
- Gizi

E. FUNGSI STATISTIKA

1. Sebagai alat bantu untuk mengumpulkan, mengolah, menganalisa dan menyimpulkan hasil.

2. Statistika dapat meningkatkan efisiensi dengan membatasi dan memastikan cara kerja dan cara pikir.
3. Statistika dapat meringkas hasil penelitian dalam bentuk yang sederhana dan mudah dipahami.
4. Statistika dapat memberikan dasar untuk melakukan interpretasi dan menarik kesimpulan.
5. Statistika dapat memberikan gambaran mengenai suatu peramalan untuk waktu yang akan datang.
6. Statistika dapat menguji/menganalisis faktor kausal dan perbedaan dari sejumlah faktor yang kompleks dan rumit.

PERANAN STATISTIK DALAM PENELITIAN

1. Menghitung besar sampel.
2. Menguji validitas dan reliabilitas instrument
3. Teknik untuk menyajikan data, antara lain tabel, grafik
4. Alat untuk menganalisis data seperti menguji hipotesis penelitian yang diajukan.

F. KEGUNAAN STATISTIK DI BIDANG KESEHATAN

Statistik Kesehatan merupakan kumpulan keterangan berbentuk angka yang berhubungan dengan masalah kesehatan masyarakat. Statistik kesehatan ini digunakan untuk :

1. Menentukan ada dan besarnya masalah kesehatan masyarakat.
2. Mengukur peristiwa penting / Vital Event yang terjadi di masyarakat.
3. menentukan prioritas masalah dan memilih alternatif pemecahan masalah kesehatan secara efisien.
4. Membuat perencanaan program kesehatan.

5. Mengadakan evaluasi pelaksanaan program kesehatan.
6. Dokumentasi untuk mengadakan perbandingan di masa mendatang.
7. Mengadakan penelitian masalah kesehatan yang belum diketahui atau menguji kebenaran suatu masalah kesehatan.
8. Memberikan penerangan tentang kesehatan kepada masyarakat / Publikasi ilmiah
9. Mengukur status kesehatan masyarakat dan mengetahui masalah kesehatan yang terdapat pada berbagai kelompok masyarakat.
10. Membandingkan status kesehatan masyarakat di satu wilayah dengan wilayah yang lain dengan rentang waktu sekarang dan masa lampau.
11. Meramalkan status kesehatan di masa mendatang.
12. Memprediksi timbulnya sampah sehingga merencanakan penyelesaiannya
13. Merumuskan perencanaan dan sistem administrasi kesehatan.

G. LANGKAH-LANGKAH METODE STATISTIK

1. Pembatasan masalah
2. Pengumpulan data
3. Pengolahan data
4. Analisa data
5. Penyajian informasi

Dalam langkah-langkah metode statistik, hal dasar yang perlu dilakukan adalah merumuskan masalah yang hendak dipecahkan dengan metode statistik.

Pembatasan masalah ini, ada dua hal dasar yakni :

1. Membatasi masalah dalam batasan *What* (apa), *who* (siapa/ subjeknya), *when* (kapan), *where* (dimana), and *how* (ukuran statistik yang dicari) dikenal (4W+1H).
2. Kemungkinan digunakan metode kuantitatif (statistik) untuk memecahkan masalah tersebut serta kemungkinan banyak data (bukan tunggal)

1.3 PENUTUP

Latihan

1. Jelaskan definisi tentang Ilmu Statistika?
2. Apakah sama “Statistika dengan Statistik” serta berikan penjelasannya?
3. Jelaskan pengertian dan tujuan Biostatistik!
4. Jelaskan peranan statistik dalam penelitian!
5. Jelaskan kegunaan statistik dalam bidang kesehatan!
6. Manfaat apakah yang dapat diambil oleh mahasiswa calon sarjana, dalam mempelajari statistika?

DAFTAR PUSTAKA

- Ating Somantri, Sambas Ali Muhidin, 2006. Aplikasi Statistika dalam Penelitian, Penerbit Pustaka Setia Bandung.
- Boediono, Wayan Koster, 2008. Teori dan Aplikasi statistika & Probabilitas, Remaja Rosdakarya, Bandung.
- Grace E.C. Korompis, 2014. Biostatistik Untuk Keperawatan, Penerbit Buku Kedokteran EGC.

2

POKOK BAHASAN PENGORGANISASIAN DATA

2.1. PENDAHULUAN

2.1.1. Deskripsi Singkat

Pengorganisasian data yang akan dibahas pada bab ini adalah jenis data, skala data, teknik pengumpulan data, teknik pengolahan data serta teknik penyajian data.

2.1.2. Relevansi

Materi yang disajikan pada pokok bahasan memberikan pemahaman dasar ilmu kesehatan masyarakat mahasiswa mampu mengidentifikasi jenis data, skala data, teknik pengumpulan data, teknik pengolahan data serta teknik penyajian data.

2.1.3. Kompetensi

a. Standar kompetensi

Pada akhir kuliah mahasiswa diharapkan dapat mengidentifikasi dan mengkomunikasikan data dalam bentuk tulisan.

b. Kompetensi Dasar

Pada akhir kuliah diharapkan mahasiswa mampu:

1. Mengidentifikasi jenis data, skala data.

2. Menjelaskan cara pengumpulan data
3. Mampu mengkomunikasikan data dalam bentuk penyajian data secara tertulis.

2.2. PENYAJIAN

2.2.1. Uraian Isi

Dalam mempelajari statistika, perlu mempelajari data yaitu segala fakta dan angka yang dapat dijadikan bahan untuk menyusun suatu informasi, sedangkan informasi adalah hasil pengolahan data yang dipakai untuk suatu keperluan. Untuk itu terdapat beberapa pengertian atau konsep yang harus diperhatikan, antara lain:

- Data
- Sumber data
- Syarat data
- Instrumen
- Sampel / Populasi

A. DATA

Data merupakan bentuk jamak sedangkan bentuk tunggalnya adalah datum, sehingga bila data kita banyak tidak perlu menyebutkan “data-data”, cukup mengucapkan data saja. Data (jamak) yaitu suatu materi/ kumpulan fakta yang dipakai untuk kepentingan suatu analisa diskusi, presentasi ilmiah atau tes statistik. Data statistik bersifat agregat (kumpulan).

JENIS DATA

Kita mengenal beberapa jenis data tergantung konteksnya. Suatu data bisa berupa angka bisa juga bukan berupa angka. Data yang berupa angka dapat disebut data kuantitatif, yang

nilainya bisa berubah-ubah. Berdasarkan nilainya dikenal dua jenis data kuantitatif, yaitu data diskrit dan data kontinu. Menurut sifatnya dikelompokkan menjadi 2 yaitu :

a. Data Kualitatif

adalah data yang dikategorikan menurut lukisan kualitas obyek yang dipelajari (atribut). Berupa data kategori.

Contohnya:

1. Karena malas belajar, Anto gagal ujian.
2. Karena suka merokok, Andi sakit jantung.

b. Data Kuantitatif

adalah data yang dinyatakan dalam bentuk numerik/angka/bilangan. Contoh : data TB yaitu 210 cm, 170 cm, 150 cm

Data tersebut dibagi menjadi 2 yaitu :

1. Data kontinyu

Yaitu data numerikal yang nilainya dapat diukur sampai sekecil-kecilnya.

Ciri-ciri :

- ✓ Diperoleh dengan cara mengukur
- ✓ Bilangan cenderung desimal
- ✓ Biasanya dinyatakan dalam nilai rata-rata

Contohnya:

- 1) Berat badan Andi 80 kg.
- 2) Tinggi badan Nina 165 cm.
- 3) Luas lapangan Parkir gedung G adalah 400 m²

2. Data Diskrit

Adalah data kategorikal yang nilainya tidak dapat diukur sekecil-kecilnya dan merupakan satu kesatuan.

Ciri-ciri :

- ✓ Diperoleh dengan cara menghitung
- ✓ Bilangan angkanya / bulat

✓ Biasanya dinyatakan dalam jumlah % atau proporsi

Contohnya:

- 1) Ibu Nani mempunyai 4 saudara.
- 2) Pak Ahmad mempunyai 250 ayam
- 3) Di Universitas Dian Nuswantoro mempunyai 5 fakultas, yaitu fakultas komputer, fakultas ekonomi, fakultas kesehatan, fakultas bahasa dan fakultas teknik industri

Menurut sumbernya, data dibedakan menjadi :

1. Data internal

Data yang dikumpulkan dari pihak internal/kalangan sendiri/ orang dalam, seperti data rekam medis, data dokter, perawat.

2. Data external

Data yang diperoleh dari pihak luar, diluar kalangan peneliti serta menggambarkan situasi kondisi yang ada di luar organisasi. Contohnya data jumlah penggunaan layanan rumah sakit.

Menurut cara memperolehnya, data ada 2 yaitu :

1. Data primer

Data utama, data yang langsung dikumpulkan oleh orang yang berkepentingan dengan data tersebut, contoh data hasil wawancara langsung pada pekerja untuk mengetahui kecelakaan kerja yang pernah dialami ditempat kerja.

2. Data sekunder

Data kedua/penunjang, data yang diperoleh atau dikumpulkan dari pihak lain, semisal data hasil penelitian pihak lain digunakan untuk data penelitian, data yang diperoleh dari studi kepustakaan.

Berdasarkan pengertian diatas, jelas data primer lebih baik daripada data sekunder, karena asal usul, kelemahan dan kelebihan data primer diketahui langsung oleh orang yang berkepentingan dengan orang tersebut, sedangkan data sekunder dikumpulkan dan diolah oleh orang lain sehingga orang luar yang memakai data tersebut tidak mengetahui asal usul, kelebihan dan kekurangan data tersebut.

Berdasarkan waktu pengumpulannya, data dibagi menjadi dua yaitu:

1. Data Cross Section

Adalah data yang menunjukkan titik waktu tertentu atau pengumpulannya dilakukan dalam waktu bersamaan. Contohnya Tingkat pengetahuan siswa SD di kota Semarang terhadap perilaku hidup bersih dan sehat pada tahun 2015.

2. Data Time Series/Berkala

Data berkala adalah data yang menggambarkan sesuatu dari waktu ke waktu atau periode secara historis. Data time series merupakan data perkembangan anak dari tahun 2010.

PENGUKURAN & SKALA DATA

Pengukuran adalah pemberian kuantifikasi pada suatu sifat yang diamati. Jadi pengukuran disini tidak hanya terbatas pada penggunaan alat ukur, seperti mistar untuk mengukur panjang, timbangan untuk mengukur berat, dll, tetapi juga mencatat umur menggunakan satuan terkecil, sehingga bila dirubah satuannya akan menjadi bilangan pecahan seperti halnya yang dihasilkan dari alat ukur.

Dalam kehidupan sehari-hari tidak semua sifat atau konsep teori dapat diukur, untuk ini perlu dilakukan “operasionalisasi” yaitu pemecahan atau penguraian sifat dalam sejumlah dimensi yang bisa diukur. Misalnya operasi-onalisasi untuk status social ekonomi dan inteligensia :

- a. “motivasi” : dimensi kehadiran
- b. “kepandaian” : skor tes IQ yang terdiri dari beberapa soal; setiap soal merupakan satu dimensi

Skala pengukuran adalah kemampuan untuk membedakan satu nilai terhadap lainnya dari suatu sifat yang diamati. Sifat yang diamati mempunyai nilai yang cenderung bervariasi. Sifat yang diamati tersebut sering disebut “variabel”

Skala pengukuran atau skala data, terbagi atas (dari skala terendah ke tertinggi):

1. Nominal

Yaitu skala yang hanya mempunyai ciri untuk membedakan skala ukur yang satu dengan skala ukur yang lain. Contoh :

- ✓ Jenis kelamin (1. laki laki 2.perempuan)
- ✓ Jenis warna (1. hijau 2. biru 3. kuning)

2. Ordinal

Yaitu skala yang selain mempunyai ciri untuk membedakan juga mempunyai ciri untuk mengurutkan pada rentangan tertentu. Misalnya rentangan dari yang paling rendah sampai yang paling tinggi, dari yang paling jelek sampai yang paling baik. Contoh :

- ✓ tingkat pendidikan : (1. SD, 2. SMP, 3. SMA)
- ✓ Sikap (sangat setuju, setuju, netral)

3. Interval

Yaitu skala yang membedakan, mempunyai arti tingkatan, mempunyai besaran/jarak/interval yang tetap antara satu data dengan yang lainnya. Contoh :

- ✓ skor
- ✓ IP

→ Data hasil pengukuran

4. Rasio

Yaitu data yang membedakan, mempunyai arti tingkatan, mempunyai besaran/jarak tertentu antar datanya, mempunyai nilai mutlak (absolute) artinya nilai '0' → kosong/ tidak ada. Contoh :

- ✓ Tinggi badan
- ✓ BB

Mengubah skala pengukuran variabel

Dalam analisis data, skala data dapat diubah dengan syarat dari skala data yang tinggi tingkatan ke skala yang lebih rendah, akan tetapi skala yang lebih rendah tidak dapat dinaikan skala datanya, semisal data rasio dapat diubah menjadi interval, ordinal bahkan nominal. Tetapi data nominal tidak dapat diubah menjadi ordinal. Berdasarkan sifat ini, maka dalam pengumpulan data usahakan dikumpulkan dalam data tertinggi.

Contoh variabel umur, bila dikumpulkan datanya, tulislah apa adanya, jangan dikelompokkan/ dikategorikan karena akan menurunkan skala data rasio menjadi paling tinggi ordinal, contoh cara menurunkan skala data sebagai berikut:

1. Menurut standar

Misalnya mengubah skala data interval menjadi skala data nominal/ dikotomi.

Skala interval	Skala nominal
Berat badan bayi lahir (gram), menurut standar dikatakan BBLR bila lahir < 2500 gram	1. BBLR 2. Non BBLR

2. Melihat gambaran univariat/ deskriptif

Misalnya variabel tinggi badan diketahui rata-rata hasil pengukuran 168cm. Dari data di atas dapat dipakai sebagai standar penentuan skala :

- 1) rendah : < 168 cm
- 2) tinggi : \geq 168 cm

PENGUMPULAN DATA

1. Wawancara

Cara untuk mengumpulkan data dengan mengadakan tatap muka secara langsung antara orang yang bertugas mengumpulkan data dengan orang yang menjadi sumber data atau obyek penelitian.

Ada dua jenis wawancara yaitu

a. wawancara berstruktur

Adalah wawancara yang sebagian besar jenis-jenis pertanyaannya telah ditentukan sebelumnya termasuk urutan yang ditanya dan materi pertanyaannya.

b. wawancara tak berstruktur

Adalah wawancara yang tidak secara ketat telah ditentukan sebelumnya mengenai jenis-jenis pertanyaan.

Kelebihannya dari wawancara:

Data yang diperlukan langsung diperoleh sehingga lebih akurat dan dapat dipertanggungjawabkan.
Data lebih akurat, lengkap, dan konsisten.

Kekurangannya dari wawancara:

Tidak dapat dilakukan dalam skala besar dan sulit memperoleh keterangan yang sifatnya pribadi, perlu dana yang besar dan pewawancara yang banyak bila data yang diperlukan besar.

2. Kuesioner

Adalah cara mengumpulkan data dengan mengirim kuesioner yang berisi sejumlah pertanyaan yang ditujukan kepada orang yang menjadi objek penelitian sehingga jawabannya tidak langsung diperoleh. Jenis, urutan dan materi pertanyaan dari kuesioner pada dasarnya hampir sama dengan wawancara.

Kelebihan dari kuesioner:

Dapat dilakukan dalam skala besar, biayanya lebih murah karena tidak perlu mengirim banyak orang, bisa memperoleh jawaban yang sifatnya pribadi.

Kekurangannya dari kuesioner:

Terdapat kemungkinan jawaban yang tidak lengkap, jawaban bisa tidak akurat hal ini terjadi bila yang menjawab tidak respondennya langsung.

3. Observasi (pengamatan)

Adalah cara mengumpulkan data dengan mengamati atau mengobservasi obyek penelitian atau peristiwa/kejadian baik berupa manusia, benda mati, maupun alam. Orang yang bertugas melakukan observasi disebut observer atau pengamat, sedangkan

alat yang dipakai untuk mengamati obyek disebut pedoman observasi.

Kelebihan dari observasi:

Data yang diperoleh lebih dapat dipercaya karena dilakukan atas pengamatan sendiri.

Kekurangan dari observasi:

Bisa terjadi kesalahan interpretasi terhadap kejadian yang diamati, bila observer berbeda dalam mengamati obyek yang sama, bisa menghasilkan kesimpulan yang berbeda.

4. Tes dan skala obyektif

Suatu cara mengumpulkan data dengan memberikan tes kepada obyek yang diteliti. Cara ini banyak dilakukan pada tes psikologi untuk mengukur karakteristik kepribadian seseorang.

5. Metode Proyektif

Adalah cara mengumpulkan data dengan mengamati atau menganalisis suatu obyek melalui ekspresi luar dari obyek tersebut dalam bentuk karya atau tulisan. Metode ini juga biasanya dipakai untuk mengetahui sikap, emosi dan kepribadian seseorang melalui karya atas tulisannya yang dibuat.

KARAKTERISTIK DATA YANG BAIK

1. Akurasi adalah data mendekati nilai yang sebenarnya
2. Presisi adalah ketelitian data sampai data sekecil kecilnya pengukuran
3. Reliabilitas yaitu stabilitas dan konsistensi data sama dengan sumber yang ada, di mana pengukuran berulang hasilnya sama
4. Validitas Eksternal

5. Adalah mengukur data yang sesungguhnya/sebenarnya dimana karakteristik data sampel harus sama dengan karakteristik populasi
6. Validitas Internal
7. Adalah menyangkut keahlian peneliti / pengumpul data serta sensitifitas alat/ instrument penelitian.

PENGOLAHAN DATA

Pengolahan data bertujuan untuk mempersiapkan data sehingga memudahkan analisa data. Prosedurnya adalah sebagai berikut :

1. Editing
yang dimaksud dengan editing adalah melihat kembali hasil pengumpulan data baik isi maupun ujud dari alat pengumpulan datanya, seperti :
 - a. Menjumlah, maksudnya menghitung jumlah lembaran daftar pertanyaan yang telah diisi, apakah sesuai jumlah yang dikehendaki. Pekerjaan ini sebaiknya dilakukan di lapangan atau di tempat yang dekat dengan sumber data, sehingga bila ada kekurangan dapat segera dilengkapi.
 - b. Koreksi, maksudnya membetulkan setiap kesalahan, kekurangan atau keraguan jawaban/hasil pengamatan yang ditemui di lapangan. Kesalahan ini adalah kesalahan yang tidak sesuai dengan tujuan pengumpulan data, semisal satuan berat badan adalah cm, maka dibetulkan menjadi kg.
2. Verifying
Langkah ini ditujukan untuk verifikasi, apakah data yang dikumpulkan dicek ulang kebenarannya, semisal melakukan kroscek ulang, seperti kroscek umur

dengan jumlah anak yang dilahirkan. Hasil akhir dari tahap ini, peneliti diyakinkan bahwa data sudah benar.

3. Coding

Pemberian kode dimaksudkan untuk mempermudah dalam pengolahan dan proses lanjutannya melalui tindakan pengklasifikasikan data. Kadang kode yang sudah dibuat perlu dirubah terutama untuk jawaban semi terbuka yang semula hanya memiliki satu kode kemudian dijadikan lebih dari satu kode, cara pengkodean demikian disebut dengan “recoding”. Selain itu juga tidak jarang diperlukan pembuatan kategori untuk data kontinyu, sehingga masing-masing mempunyai kode sendiri-sendiri atau data kategori itu akan digunakan sebagai penyajian data secara deskriptif. Pada langkah ini diperlukan disesuaikan dengan skala data yang ada. Perlu diingat dan dibuat selalu daftar kode berikut keterangannya, agar dalam uraian/ bahasan tentang hasil pengamatan tidak salah menerjemahkannya.

4. Classifying

Tahapan ini merupakan tahapan untuk mengelompokan data menurut kategori tertentu. Kategori ditetapkan oleh peneliti, sehingga data yang klasifikasinya memenuhi dapat dikelompokan dan dihitung. Apabila tahapan ini sudah dilakukan, maka akan memudahkan pada tahapan tabulating.

5. Tabulating

Tabulating adalah mengorganisir data sedemikian rupa sehingga mudah untuk dijumlah, disusun, dan disajikan dalam bentuk tabel atau grafik. Sesuai dengan fasilitas yang digunakan, penyusunan data melalui komputer akan memberikan hasil yang jauh lebih baik. Sebelum

era komputerisasi, pembuatan tabel yang berisi data secara keseluruhan dikenal dengan nama “*Master Table*”, yang berisi tabulasi data secara keseluruhan, menjadi tabel utama. Dari tabel ini kemudian dapat dibuat tabel berikutnya, semisal tabel silang.

PENYAJIAN DATA

Penyajian data adalah pemaparan data hasil pengamatan yang telah disusun secara teratur, sehingga hasil pengamatan tersebut dapat dipahami dengan baik. Penyajian dilakukan apabila hendak mempresentasikan hasil analisis data yang telah dianalisis. Prinsip utama penyajian adalah pada diterimanya informasi hasil penyajian oleh pihak, sebagai tujuan penyajian. Dengan demikian penyajian yang terbaik adalah penyajian yang dimengerti oleh penerima penyajian, dan sesuai dengan kebutuhan penerima penyajian.

Ada tiga jenis penyajian yaitu :

1. Tulisan (textular presentation)
Narasi/ tulisan untuk data kecil dan kesimpulan yang sederhana
2. Tabel (table presentation)
Dalam bentuk tabel untuk menyajikan data dalam beberapa variabel sekaligus
3. Grafik (graphical presentation)
Disajikan dalam bentuk grafik (gambaran) berupa batang, histogram, pie, line, kurtogram.

Tulisan (textular presentation)

Bentuk penyajian dalam bentuk tulisan ini membuat keterangan tentang prosedur, hasil dan kesimpulan secara

garis besar, sehingga tidak banyak diperoleh gambaran secara statistik.

Contoh :

Jumlah penderita yang dirawat di RS.SEHAT SLALU pada tahun 2013 dirinci sebagai berikut :

Bagian penyakit Dalam	sebanyak	150 orang
Bagian Anak	sebanyak	240 orang

.....

Jumlah		1080 orang
--------	--	------------

TABEL (table presentation)

Penyajian dalam bentuk ini ternyata yang paling lazim digunakan, agar dapat diperoleh gambaran yang lebih terperinci tentang suatu variabel disamping nantinya bisa pula digunakan untuk perbandingan-perbandingan. Pembuatan tabel frekuensi (variabel tunggal) sering juga dilakukan sebagai dasar pembuatan grafik.

Pembagian tabel :

1. Menurut fungsinya :
 - tabel sinopsis
 - tabel induk
 - tabel teks
 - tabel kontigensi
2. Menurut judul baris :
 - menurut abjad, menurut geografis
 - perkembangan waktu, besarnya angka

Kelengkapan tabel :

1. nomor tabel
2. judul
3. catatan pendahuluan

4. badan tabel, terdiri dari :judul kolom, judul baris, judul kompartement, jalur-jalur, baris-baris dan sel
5. catatan kaki
6. sumber data

Contoh tabel:

Tabel. Produktivitas pertanian menurut dosis pestisida yang digunakan di desa Sumber Rejo kecamatan Ngablak Magelang 2009.

Dosis Pestisida	Produktivitas				Total
	Rugi	%	Untung	%	
Tidak sesuai anjuran	25	42	35	58	60
Sesuai anjuran	2	25	6	75	8
Total	27		41		68

Hasil $X^2 = 0,271$ dan $p = 0,603$

RP (95% CI) = 2,14 (0,399-11,504)

GRAFIK (graphical presentation)

Kegunaan grafik antara lain :

1. Membandingkan beberapa variabel atau variabel yang sama menurut waktu atau tempat yang berbeda
2. Meramalkan suatu keadaan di masa yang akan datang
3. Mengetahui hubungan satu atau lebih variabel
4. Memberikan penerangan kepada masyarakat

Keuntungan garfik

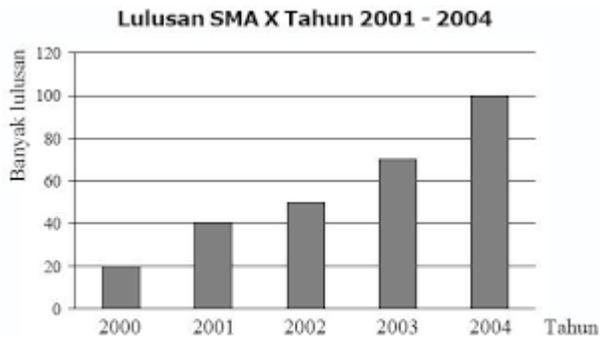
1. Pengamat lebih mudah memahami masalah yang disajikan
2. Tampak lebih menarik dibandingkan tabel
3. Hal-hal yang tidak jelas dalam tabel dapat diperjelas lagi dalam bentuk grafik

Penyajian dengan grafik harus melihat kebutuhan penerima/ sesuai tujuan presentasi, diantaranya :

1. Diagram Batang

Diagram batang sangat cocok untuk menyajikan data yang berbentuk kategori atau atribut, dan data tahunan yang tahunnya tidak terlalu banyak. Untuk menggambar diagram batang diperlukan sumbu tegak dan sumbu datar yang berpotongan tegak lurus. Sumbu tegak maupun sumbu datar dibagi menjadi beberapa skala bagian yang sama. Pada bagian bawah sumbu datar dituliskan atribut atau waktu dan pada sumbu tegak dituliskan kuantum atau nilai data.

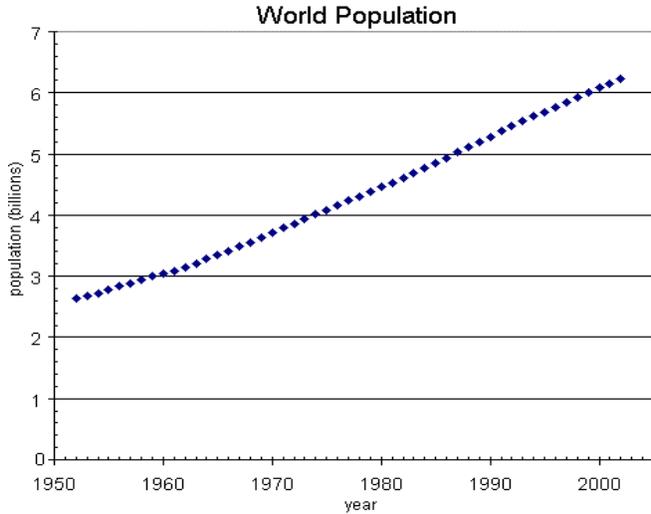
Contohnya :



2. Diagram garis

Diagram garis sangat cocok untuk menyajikan data yang berbentuk serba terus atau berkesinambungan, semisal pada data yang disajikan menurut waktu. Untuk menggambar diagram garis diperlukan sumbu tegak dan sumbu datar yang berpotongan tegak lurus. Sumbu tegak maupun sumbu datar dibagi menjadi beberapa bagian yang sama yaitu bagian bawah sumbu datar dituliskan urutan atau waktu dan pada sumbu tegak dituliskan kuantum atau nilai data.

Contohnya sebagai berikut :



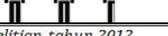
source: U.S. Bureau of the Census, International Data Base (via GeoHive)

3. Diagram Lambang

Diagram Lambang sangat cocok untuk menyajikan data kasar sesuatu hal dan sebagai alat visual bagi orang awam. Setiap satuan yang dijadikan lambang disesuaikan dengan macam datanya. Misalnya untuk data jumlah manusia dibuatkan gambar orang. Satu gambar orang menyatakan sekian jiwa tergantung kebutuhannya. Kelemahannya ialah jika data yang dilaporkan tidak penuh (bulat) sehingga lambangnya pun menjadi tidak utuh.

Contohnya:

GAMBAR 1.3 DIAGRAM LAMBANG

RW 012 (RT)	JUMLAH KEPALA KELUARGA (LAMBANG)
RT 001: 38	
RT 002: 34	
RT 003: 32	
RT 004: 26	

Sumber : Hasil penelitian tahun 2012

Keterangan:



Mewakili 10 orang

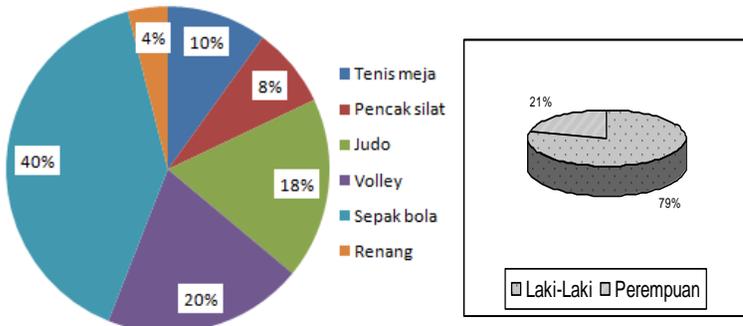
4. Diagram Lingkaran atau Diagram Pastel

Diagram lingkaran sangat cocok untuk menyajikan data yang berbentuk kategori atau atribut dalam presentase. Untuk membuat diagram lingkaran, maka lingkaran dibagi-bagi menjadi beberapa sektor. Setiap sektor melukiskan kategori data yang terlebih dahulu diubah ke dalam derajat dengan menggunakan busur derajat.

Jika datanya 25%, maka gambar derajatnya adalah :
 $25/100 \times 360 = 90$

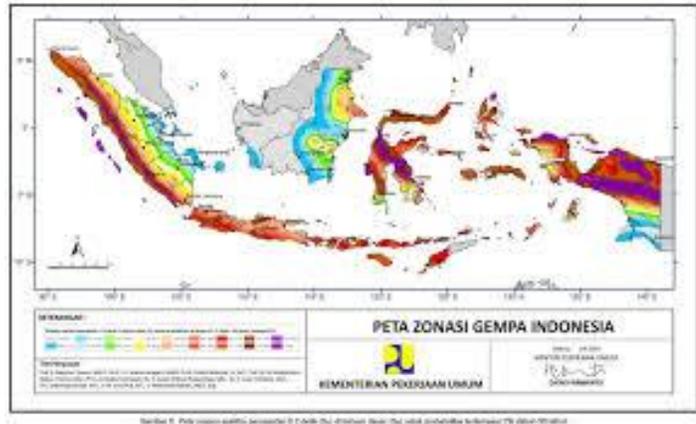
Bila diagram lingkaran ini digambarkan perspektifnya menjadi gambar tiga dimensi, maka diagramnya disebut diagram pastel.

Contohnya :



5. Diagram Peta (Kartogram)

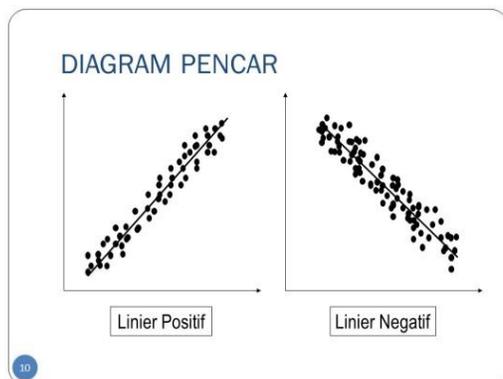
Diagram peta sangat cocok untuk menyajikan data yang ada hubungannya dengan tempat kejadian.



6. Diagram Pencar (Titik)

Diagram pencar sangat cocok untuk menyajikan data yang terdiri atas dua variabel. Diagramnya dibuat dalam sistem koordinat. Diagram pencar ini berfungsi pula untuk menentukan apakah suatu data linier.

Contohnya :

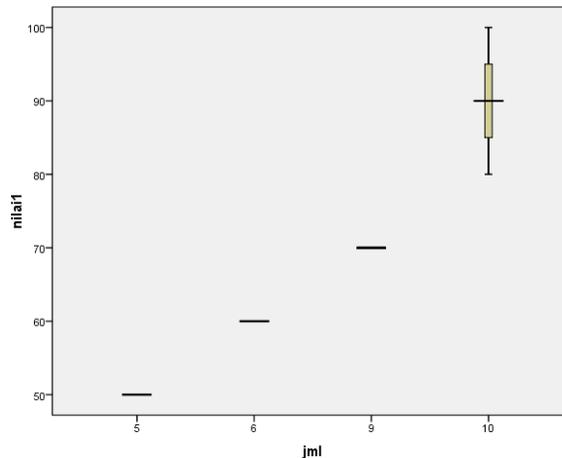


7. Box and Whisker Plot

Digunakan untuk menyajikan data numerik dan diagram tersebut dipakai untuk membandingkan beberapa pengamatan. Kotak (box) terdiri atas:

- Garis tengah adalah nilai kuartil dua Q2 atau median
- Garis bawah adalah nilai kuartil satu Q1
- Garis atas kotak adalah nilai kuartil tiga Q3

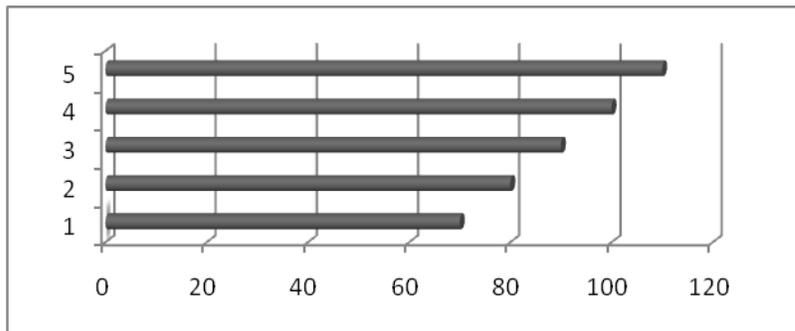
Tali (*whisker*) batas bawah adalah nilai batas yang tidak lebih perbedaannya dengan Q1 sebanyak $1 \frac{1}{2} \times (Q3-Q2)$ atau perbedaan interkuartil, sedangkan batas atas adalah nilai yang paling jauh dan tidak lebih dari $1 \frac{1}{2} \times (Q3-Q1)$.



8. Pareto Chart

Pareto tidak berbeda dengan diagram batang yang disusun dengan susunan tinggi rendahnya batang sehingga dengan mudah dapat diinterpretasi.

Contohnya:



2.3. PENUTUP

Latihan

1. Uraikan klasifikasi data beserta contohnya!
2. Jelaskan skala data dan berikan contohnya!
3. Apakah skala data dapat diubah, bila jawabannya “ya” silahkan berikan contoh merubahnya serta bila jawaban “tidak” berikan alasannya?
4. Jelaskan beberapa metode dalam pengumpulan data!
5. Analisislah masing-masing metode dalam pengumpulan data kelemahan & kelebihanannya!
6. Jelaskan beberapa jenis penyajian data!
7. Sebutkan masing-masing kelebihan maupun kelemahan dari penyajian data!
8. Jelaskan kelebihan dan kekurangan dari beberapa macam diagram!
9. Buatlah contoh penyajian data dengan grafik!
10. Jelaskan prosedur pengolahan data!

DAFTAR PUSTAKA

- Ating Somantri, Sambas Ali Muhidin, 2006. Aplikasi Statistika dalam Penelitian, Penerbit Pustaka Setia Bandung.
- Boediono, Wayan Koster, 2008. Teori dan Aplikasi statistika & Probabilitas, Remaja Rosdakarya, Bandung.
- Grace E.C. Korompis, 2014. Biostatistik Untuk Keperawatan, Penerbit Buku Kedokteran EGC.
- Husaini usman dkk, 2003. Pengantar Statistika, Bumi Aksara, Jakarta.

3

POKOK BAHASAN DISTRIBUSI FREKUENSI

3.1. PENDAHULUAN

3.1.1. Deskripsi Singkat

Distribusi frekuensi pada bab ini akan membahas pengertian distribusi frekuensi dan penerapannya.

3.1.2. Relevansi

Materi yang disajikan pada pokok bahasan memberikan pemahaman mahasiswa dalam menyusun distribusi frekuensi.

3.1.3. Kompetensi

a. Standar Kompetensi

Pada akhir kuliah mahasiswa diharapkan dapat menyajikan data dalam bentuk distribusi frekuensi.

b. Kompetensi Dasar

Pada akhir kuliah diharapkan mahasiswa mampu menjelaskan pengertian distribusi frekuensi serta dapat menerapkan dalam contoh soal.

3.2. PENYAJIAN

3.2.1. Uraian Isi

Distribusi frekuensi adalah suatu bentuk penyusunan yang teratur mengenai suatu rangkaian data, dengan menggolongkan besar dan kecilnya angka-angka yang bervariasi ke dalam kelas-kelas tertentu. Suatu rangkaian data disusun ke dalam suatu distribusi frekuensi untuk memperoleh gambaran yang sederhana, jelas, dan sistematis mengenai rangkaian data tersebut.

Data yang dikelompokkan mempunyai kelebihanannya mempunyai gambaran menyeluruh secara jelas mengenai data yang kita miliki, tetapi mempunyai kekurangan yaitu rincian data atau informasi awal menjadi hilang sehingga data berkelompok menjadi semu atau tidak nyata.

Distribusi frekuensi dibagi menjadi 2 yaitu :

1. Distribusi frekuensi untuk data tunggal (ungrouped data)

Contoh: distribusi frekuensi untuk nilai mata kuliah X mahasiswa semester Y fakultas Z tahun 2014

Nilai X	Frekuensi (f)
80	4
70	23
60	28
50	16

2. Distribusi frekuensi untuk data berkelompok (grouped data)

Pada umumnya apabila membicarakan mengenai distribusi frekuensi, maka yang dimaksud adalah distribusi frekuensi untuk data berkelompok.

Contoh tabel Distribusi frekuensi nilai biostatistik mahasiswa kesehatan Udinus

Klas Interval	Frekuensi	Persentase (%)
51 – 60	15	15,0
61 – 70	25	25,0
71 – 80	40	40,0
81 – 90	15	15,0
91 – 100	5	5,0
Jumlah	100	100,0

Beberapa istilah dalam menyusun distribusi frekuensi, antara lain:

a. TEPI KELAS/LIMIT KELAS

Tepi kelas ada 2 yaitu

- Tepi kelas atas biasanya terletak di deret sebelah kanan contoh: dari tabel diatas adalah 60,70,80,90,100.
- Tepi kelas bawah biasanya terletak disebelah kiri, contohnya pada tabel diatas adalah 51,61,71,81
- Batas kelas selalu dinyatakan satu desimal lebih banyak/lebih sedikit daripada data atau pengamatan aslinya. Hal ini dimaksud untuk menjamin bahwa tidak akan ada data yang jatuh atau terletak tepat pada pada batas kelas.

Batas kelas atas dari tabel diatas adalah 60,5; 70,5; 80,5; 90,5;100,5

Batas kelas bawah dari tabel diatas adalah 50,5; 60,5; 70,5; 80,5; 90,5

b. Lebar Kelas

- Lebar kelas adalah jumlah nilai-nilai variabel dalam tiap-tiap kelas.

- Lebar kelas batas atas nyata dikurang batas bawah nyata dari kelas yg bersangkutan
contoh : $60,5-50,5=10$

jadi lebar kelas tabel diatas adalah 10

c. Titik Tengah

- Adalah angka atau nilai variabel yang terdapat ditengah-tengah interval kelas.
- Nilai tengah antara batas bawah kelas dengan batas atas kelas tersebut.

$$\text{Nilai tengah kelas} = \frac{\text{Batas bawah kelas} + \text{Batas atas kelas}}{2}$$

Contoh:

13, 14, 15 titik tengahnya adalah 14

20,21,22,23 titik tengahnya adalah separo dari jumlah angka tengah yaitu $(21+22) \times \frac{1}{2} = 21,5$

Nilai tengah dari tabel diatas adalah $(50,5 + 60,5)/2 = 55,5$ dan seterusnya

d. Jumlah Interval

- Adalah banyaknya interval yang digunakan dalam penyusunan distribusi

Contoh: dari tabel diatas ada 5

51 – 60

61 – 70

71 – 80

81 – 90

91 – 100

e. Jarak Pengukuran

- Adalah angka tertinggi dari pengukuran dikurangi dengan angka terendah

- R adalah batas nyata atas (upper real limit) dari nilai variabel yang tertinggi dikurangi dengan batas nyata bawah (lower real limit) dari nilai variabel yang terendah

Prosedur pembuatan tabel distribusi dengan klas interval

Menentukan jumlah klas interval

Ada 3 cara :

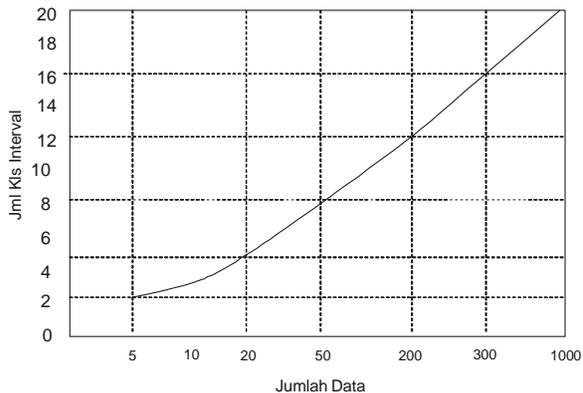
1). Berdasarkan pengalaman

Dilakukan oleh orang yang sudah biasa/pengalaman dalam pengolahan data.

2). Dengan Grafik

misal :

- Bila jumlah data 50, maka jml klas intervalnya 8
- Bila jumlah data 200, maka jml klas intervalnya 12



3). Rumus Sturges :

1. Menentukan Jumlah klas interval

$$k = 1 + 3,3 \log n$$

k = jumlah klas interval

n = jumlah data

log = logaritma

2. Menentukan rentang data

Rentang data adalah selisih nilai data terbesar dengan nilai data terkecil.

3. Menghitung panjang klas interval

Panjang klas interval = rentang : jumlah klas interval

Menyusun klas interval dalam distribusi frekuensi.

Supaya komunikatif tidak selalu dimulai dari nilai yang terkecil.

Prosedur pembuatan tabel distribusi dengan klas interval

Contoh :

60	75	65	80	56	70	85	90	65	60
76	58	65	70	80	68	66	78	74	65
50	60	56	76	84	68	90	86	68	66
44	60	65	58	72	64	50	58	82	76
80	86	64	78	50	64	72	84	42	80
66	80	64	76	82	68	84	90	50	62

Hasil perhitungan

- Jumlah kelas

$$k=1 + 3.3 \log n$$

$$k= 1 + 3.3 \log 60$$

$$k= 7$$

- Rentang data $90.5-41.5 = 49$

- Panjang kelas interval = rentang kelas/jml kls

$$49/7=7$$

Hasil perhitungan

Interval Kls	Frekuensi
42-48	2
49-55	4
56-62	10
63-69	16
70-76	10
77-83	9
84-90	9
Jumlah	60

Contoh penyelesaian distribusi frekuensi, bila diketahui distribusi data sebagai berikut :

DISTRIBUSI DATA				
88	55	48	36	61
42	58	55	69	63
25	47	78	61	54
67	40	51	59	68
83	59	13	72	57
53	82	69	60	35
42	55	34	49	45
44	59	46	71	86
51	75	44	66	53
60	41	29	56	57
27	62	44	85	61
53	58	26	77	68
52	49	45	54	41
45	59	44	68	73
51	85	46	55	67

Langkah-langkah pembentukan distribusi frekuensi :

1. tentukan jumlah kelas (k) = $1 + 3,3 \log N$, dengan N adalah jumlah data
 $\log 75 = 1,88$
 - a. $k = 1 + 3,3 \log n$

$$\begin{aligned}
 &= 1 + 3,3 \log 75 \\
 &= 1 + 3,3 \cdot 1,88 \\
 &= 7,204 \rightarrow 7 \text{ atau } 8
 \end{aligned}$$

2. tentukan range = nilai data terbesar – nilai data terkecil
 $88 - 13 = 75$

3. tentukan lebar kelas interval (I) = R/K
 $75 : 8 = 9,375 \rightarrow 9 \text{ atau } 10$

Untuk menyusun interval kelas = tepi kelas bawah + panjang kelas - 1

$$= 10 + (10-1) = 19, \text{ dst}$$

4. menyusun tabel tally

Kelas interval	Tally	Frekuensi
10-19	I	1
20-29	IIII	4
30-39	III	3
40-49	IIII IIII IIII II	18
50-59	IIII IIII IIII IIII II	22
60-69	IIII IIII IIII	15
70-79	IIII I	6
80-89	IIII I	6
Jumlah		75

5. Menyempurnakan tabel distribusi frekuensi

Kelas interval	Frekuensi
10-19	1
20-29	4
30-39	3
40-49	18
50-59	22
60-69	15
70-79	6
80-89	6
Jumlah	75

Distribusi Frekuensi Relatif

Pada beberapa jenis analisa statistik, menggunakan distribusi frekuensi relatif yaitu disajikan dalam bentuk persentase.

Contoh :

Kelas interval	Frekuensi relatif (%)
10-19	1,33
20-29	5,34
30-39	4
40-49	24
50-59	29,33
60-69	20
70-79	8
80-89	8

- a. $1/75 \times 100\% = 1,33\%$
- b. $4/75 \times 100\% = 5,33\%$
- c. $3/75 \times 100\% = 4\%$
- d. $18/75 \times 100\% = 24\%$
- e. $22/75 \times 100\% = 29,33\%$
- f. $15/75 \times 100\% = 20\%$
- g. $6/75 \times 100\% = 8\%$

Distribusi Frekuensi Kumulatif

Pada beberapa jenis analisa statistik, menggunakan distribusi frekuensi kumulatif. Terdapat 2 jenis distribusi frekuensi kumulatif yaitu :

Distribusi kumulatif “kurang dari” dan Distribusi kumulatif “lebih dari/ lebih”. Pada Distribusi kumulatif “kurang dari” dan Distribusi kumulatif “lebih dari/ atau lebih”, penentuan

frekuensi tiap-tiap kelas dengan menambah atau mengurangi secara berturut-turut kelas-kelas sebelumnya.

Distribusi frekuensi kumulatif Kurang dari

Kelas interval	Frekuensi kumulatif
Kurang dari 10	0
Kurang dari 20	1
Kurang dari 30	5
Kurang dari 40	8
Kurang dari 50	26
Kurang dari 60	48
Kurang dari 70	63
Kurang dari 80	69
Kurang dari 90	75

Lebih dari

Kelas interval	Frekuensi kumulatif
Kurang dari 10	75
Kurang dari 20	74
Kurang dari 30	70
Kurang dari 40	67
Kurang dari 50	49
Kurang dari 60	27
Kurang dari 70	12
Kurang dari 80	6
Kurang dari 90	0

Penyajian Grafik Distribusi Frekuensi

1. Histogram frekuensi

Menggambarkan beda antara frekuensi kelas-kelas dalam sebuah distribusi frekuensi, dalam bentuk bar/persegi panjang.

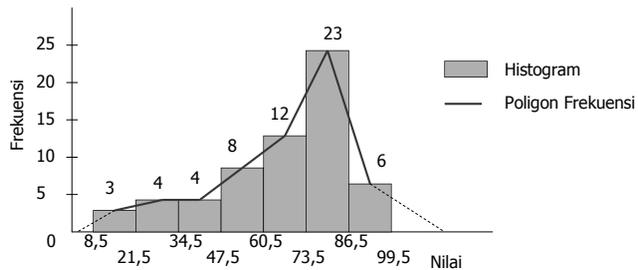
Sumbu X : interval kelas (menggunakan batas kelas)

Sumbu Y : frekuensi masing-masing kelas

2. Polygon frekuensi
 digunakan untuk membandingkan antara 2 atau beberapa distribusi frekuensi
 cara : dengan menentukan titik tengah bagi tiap puncak persegi panjang dalam histogram, kemudian menghubungkannya dengan garis linier.
3. Kurva Ogive
 menggambarkan frekuensi “kurang dari” dan “atau lebih” dari masing-masing kelas pada ordinat yang dihubungkan
 Sumbu X : interval kelas (menggunakan batas kelas)
 Sumbu Y : frekuensi “kurang dari” dan “lebih” masing-masing kelas

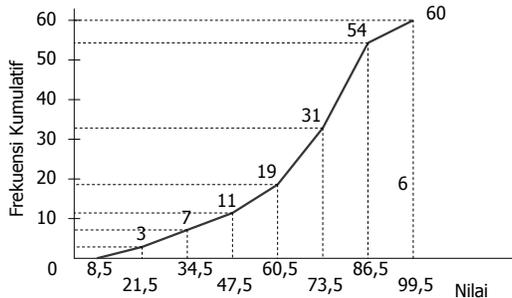
Contoh Histogram & Poligon

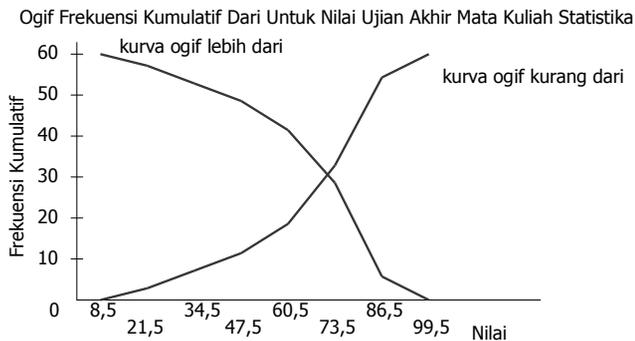
Histogram dan Poligon Frekuensi Nilai Ujian Akhir Mata Kuliah Statistika



Contoh Ogive

Ogif Frekuensi Kumulatif Kurang Dari Untuk Nilai Ujian Akhir Mata Kuliah Statistika





3.3. PENUTUP

Latihan

1. Jelaskan yang dimaksud dengan limit kelas, batas kelas, nilai tengah kelas, range, frekuensi kumulatif!
2. Sebutkan kelebihan dan kelemahan penyajian data dengan distribusi frekuensi!
3. Dari tabel dibawah ini tentukan
 - a. Buatlah tabel distribusi frekuensi yang berisi kelas interval, batas kelas, nilai tengah, frekuensi dan frekuensi relatif
 - b. Buatlah tabel distribusi frekuensi kumulatif lebih dari dan kurang dari!
 - c. Buatlah histogram, polygon dan kurva ogive!

Data hasil ujian akhir Mata Kuliah Statistika dari 60 orang mahasiswa

25	60	79	32	57	74	52	70	82	36
75	77	81	95	41	65	92	85	55	66
52	10	64	75	78	25	80	98	81	67
41	71	83	54	64	72	88	62	74	45
60	78	89	76	84	48	84	90	15	79
35	67	17	82	69	74	63	80	85	61

DAFTAR PUSTAKA

- Ating Somantri, Sambas Ali Muhidin, 2006. Aplikasi Statistika dalam Penelitian, Penerbit Pustaka Setia Bandung.
- Boediono, Wayan Koster, 2008. Teori dan Aplikasi statistika & Probabilitas, Remaja Rosdakarya, Bandung.
- Grace E.C. Korompis, 2014. Biostatistik Untuk Keperawatan, Penerbit Buku Kedokteran EGC.
- Husaini usman dkk, 2003. Pengantar Statistika, Bumi Aksara, Jakarta.

4

POKOK BAHASAN UKURAN PEMUSATAN DATA

4.1. PENDAHULUAN

4.1.1. Deskripsi Singkat

Pada bab ukuran pemusatan data yang akan dibahas adalah tentang menerapkan rumus ukuran tendensi sentral yaitu mean, median, modus, kuartil, desil dan persentil.

4.1.2. Relevansi

Materi yang disajikan pada bab ini meliputi: menerapkan rumus mean, median, modus, kuartil, desil dan persentil.

4.1.3. Kompetensi

a. Standar Kompetensi

Pada akhir kuliah mahasiswa diharapkan dapat menjelaskan dan menerapkan rumus ukuran tendensi sentral.

b. Kompetensi Dasar

Pada akhir kuliah diharapkan mahasiswa mampu:

1. Menggunakan rumus ukuran tendensi sentral (mean, median, modus baik untuk data tunggal maupun kelompok).
2. Menghitung kuartil dan desil pada contoh kasus data.

4.2. PENYAJIAN

4.2.1. Uraian Isi

Ukuran pemusatan dan letak data masih merupakan bagian dari statistika deskriptif. Ukuran pemusatan data yang akan kita pelajari adalah rata-rata hitung, median, modus. Sedangkan ukuran letak data yaitu kuartil, desil dan persentil. Ukuran pemusatan data disebut juga rata-rata menunjukkan dimana suatu data memusat atau suatu kumpulan pengamatan memusat (mengelompok).

a. Nilai rata-rata (mean)

lalah suatu bilangan yang dapat dipakai sebagai wakil dari sekelompok data.

$$\text{Rata - rata hitung} = \frac{\text{Jumlah semua nilai data}}{\text{Banyaknya nilai data}}$$

1) Rata-rata hitung

- Data yang belum berkelompok (ungrouped data)

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

Dimana

xi : nilai data

N : jumlah data

Contoh : dari data 9, 5, 7, 3, 4, 2

Nilai rata-rata $x = 30/5 = 5$

- Data yang berkelompok

Rata-rata hitung dalam tabel distribusi frekuensi

Modal	Nilai tengah	Frekuensi	fX
112-120	116	4	464
121-129	125	5	625
130-138	134	8	1072
139-147	143	12	1716
148-156	152	5	760
157-165	161	4	644
166-174	170	2	340
		$\Sigma f = 40$	$\Sigma fX = 5621$

Nilai rata-rata dari tabel diatas adalah:

$$\bar{X} = \frac{\Sigma fX}{\Sigma f} = \frac{5621}{40} = 140,25$$

Rata-rata hitung dengan menggunakan kode (U)

$$\bar{X} = X_0 + c \left[\frac{\Sigma fU}{\Sigma f} \right]$$

Dimana:

X_0 : nilai tengah kelas yang berhimpit dengan nilai $U=0$

c : lebar kelas

U : kode kelas

Modal	Nilai tengah	U	Frekuensi	fU
112-120	116	-3	4	-12
121-129	125	-2	5	-10
130-138	134	-1	8	-8
139-147	143	0	12	0
148-156	152	1	5	5
157-165	161	2	4	8
166-174	170	3	2	6
			$\Sigma f = 40$	$\Sigma fU = -11$

Untuk data diatas, diperoleh $X_o = 143$, $\Sigma f = 40$, $\Sigma fU = -11$ dan lebar kelas $c = 9$

$$\begin{aligned} \bar{X} &= X_o + c \left[\frac{\Sigma f U}{\Sigma f} \right] \\ &= 143 + 9 \left\{ \frac{-11}{40} \right\} = 143 - 2,475 = 140,525 \end{aligned}$$

2) Rata-rata yang ditimbang

Untuk menghitung nilai rata-rata dari suatu nilai rata-rata atau angka index relatif lainnya.

Contoh :

Sampel A, $N = 8$, $x = 45$

Sampel B, $N = 3$, $x = 20$

Sampel C, $N = 6$, $x = 30$

$$x_{ABC} = \frac{N_A \cdot x_A + N_B \cdot x_B + N_C \cdot x_C}{N_A + N_B + N_C} = \frac{8 \cdot 45 + 3 \cdot 20 + 6 \cdot 30}{8 + 3 + 6}$$

Sifat rata-rata hitung (mean)

1. dipengaruhi oleh bilangan ekstrem yang terlalu besar atau terlalu kecil sehingga tidak menggambarkan ukuran gejala pusat
2. dipengaruhi semua nilai pengamatan sehingga bila salah satu nilai tidak diketahui maka rata-rata hitung tidak dapat diketahui
3. paling banyak dikenal dan mudah menghitungnya

b. Median (Md)

ialah suatu ukuran gejala pusat yang menunjukkan letak dan membagi sekumpulan bilangan menjadi 2 sehingga separo bilangan \geq median dan separo bilangan \leq median

1) Ungrouped data

Posisi Me = $N + \frac{1}{2}$, dengan data harus dalam bentuk array / urutan

Contoh :

2 4 100 7 3 2 7 ---- array 2 2 3 4 7 7 100

Posisi Me = $8/2 = 4$, nilai Me = 4

2) Grouped data

$$\text{Med} = L_0 + c \left(\frac{\frac{n}{2} - F}{f} \right)$$

L_0 = batas bawah kelas median

F = jumlah frekuensi semua kelas sebelum kelas yang mengandung median

f = frekuensi kelas median

Contoh:

Interval Kelas	Frekuensi
9-21	3
22-34	4
35-47	4
48-60	8
61-73	12
74-86	23
87-99	6
$\Sigma f = 60$	

Letak median ada pada data ke 30 yaitu pada interval 61-73 sehingga

$$Lo = 60,5$$

$$F = 19$$

$$f = 12$$

$$c = 13$$

$$Med = 60,5 + 13 \left(\frac{\frac{60}{2} - 19}{12} \right) = 72,42$$

c. Modus

ialah nilai yang paling banyak terdapat dalam sekumpulan bilangan

1) Ungrouped data

162 157 171 143 154 : tidak ada modus

159 162 163 162 156 : modus 162

157 169 149 169 157 : bimodal 169 dan 157

2) Grouped data

$$Mod = Lo + c \left\{ \frac{b1}{b1 + b2} \right\}$$

Dimana :

Mod : modus

Lo : batas bawah kelas modus yaitu kelas interval dengan frekuensi terbanyak

C : lebar kelas modus

b1 : frekuensi kelas modus-frekuensi kelas terdekat sebelumnya

b2 : frekuensi kelas modus-frekuensi kelas terdekat sesudahnya

Modal	Frekuensi
112-120	4
121-129	5
130-138	8
139-147	12
148-156	5
157-165	4
166-174	2
Σf = 40	

Pada tabel distribusi frekuensi tersebut kelas interval 139-147 mempunyai frekuensi 12 dan merupakan frekuensi terbesar, sehingga modulusnya terletak pada kelas 139-147. Dengan demikian $Lo = 138,5$ $c = 9$, $b_1 = 12 - 8 = 4$, $b_2 = 12 - 5 = 7$ jadi modulusnya adalah

$$Mod = 138,5 + 9 \left\{ \frac{4}{4 + 7} \right\} = 138,5 + 3,27 = 141,77$$

Sifat modulus : tidak terpengaruh oleh nilai ekstrem

d. Hubungan antara Mean, Median, dan Modus

Ada 3 kemungkinan kesimetrian kurva distribusi data :

- 1) Jika nilai ketiganya hampir sama maka kurva mendekati simetri.
- 2) Jika $Mod < Med < \text{rata-rata hitung}$, maka kurva miring ke kanan.
- 3) Jika $\text{rata-rata hitung} < Med < Mod$, maka kurva miring ke kiri

e. Penggunaan Mean, Median, dan Modus

- 1) Mean
 - ✓ Bila ingin mendapatkan ukuran tendensi tengah yang stabil
 - ✓ Bila akan menghitung nilai penyebaran
 - ✓ Bila distribusinya simetris
- 2) Median
 - ✓ Bila ingin mendapatkan ukuran tendensi tengah yang lebih mudah menghitungnya
 - ✓ Bila ada nilai ekstrem yang terlalu besar / kecil

3) Modus

- ✓ Bila hanya ingin mendapatkan ukuran tendensi tengah yang kasar.
- ✓ Bila ingin mendapatkan nilai paling umum dari suatu kelompok data.

f. Kuartil

Bilangan pembagiannya ada 3 masing-masing disebut kuartil yaitu kuartil pertama (Q1)/kuartil bawah, kuartil kedua (Q2)/tengah, kuartil ketiga (Q3)/kuartil atas.

Nilai kuartil ke-i, ditentukan dengan rumus sebagai berikut:

Untuk data tidak berkelompok :

$$Q_i = \text{Nilai yang ke } - \frac{i(n+1)}{4} \quad i=1,2,3$$

Untuk data berkelompok :

$$Q_i = L_0 + c \left\{ \frac{\left(\frac{in}{4}\right) - F}{f} \right\} \quad i=1,2,3$$

Dimana

L_0 = batas bawah kelas kuartil

C = lebar kelas

F = jumlah frekuensi semua kelas sebelum kelas kuartil Q_1

f = frekuensi kelas kuartil Q_1

contoh:

Tentukan nilai kuartil Q_1 , Q_2 , Q_3 dari 10 nilai ujian biostatistik mahasiswa kesehatan, sebagai berikut: 50, 40, 45, 60, 75, 80, 80, 78, 90, 100

Langkah pertama yang dilakukan adalah mengurutkan data:

40, 45, 50, 60, 75, 78, 80, 80, 90, 100

Menentukan

$$Q_i = \text{nilai ke } \frac{i(n+1)}{4} \quad \text{dimana } n=10$$

Maka nilai kuartil Q_1 , Q_2 , Q_3

$$\begin{aligned}
 Q_1 &= \text{nilai ke } \frac{1(10+1)}{4} = \text{nilai ke } 11/4 = \text{nilai ke } 2,75 \\
 &= \text{Nilai ke-2} + 0,75 (\text{nilai ke-3} - \text{nilai ke-2}) \\
 &= 45 + 0,75 (50-45) \\
 &= 45 + 0,75 (5) \\
 &= 45 + 3,75 \\
 &= 48,75
 \end{aligned}$$

$$\begin{aligned}
 Q_2 &= \text{nilai ke- } (2(10+1))/4 = 21/4 = 5 \frac{1}{4} \\
 &= \text{nilai ke-5} + \frac{1}{4}(\text{nilai ke-6} - \text{nilai ke-5}) \\
 &= 75 + \frac{1}{4}(78-75) \\
 &= 75 + 0,75 \\
 &= 75,75
 \end{aligned}$$

$$\begin{aligned}
 Q_3 &= \text{nilai ke- } (3(10+1))/4 = 31/4 = 7 \frac{3}{4} \\
 &= \text{nilai ke } 7 + 0,75 (\text{nilai } 8 - \text{nilai } 7) \\
 &= 80 + 0,75 (80-80) \\
 &= 80 + 0 \\
 &= \mathbf{80}
 \end{aligned}$$

Contoh data kelompok

Contoh :

Interval Kelas	Nilai Tengah (X)	Frekuensi
9-21	15	3
22-34	28	4
35-47	41	4
48-60	54	8
61-73	67	12
74-86	80	23
87-99	93	6
		$\Sigma f = 60$

Q_1 membagi data menjadi 25 %
 Q_2 membagi data menjadi 50 %
 Q_3 membagi data menjadi 75 %

Sehingga :

Q_1 terletak pada 48-60
 Q_2 terletak pada 61-73
 Q_3 terletak pada 74-86

Untuk Q_1 , maka :

$$Q_1 = 47,5 + 13 \left(\frac{1,60 - 11}{4} \right) = 54$$

Untuk Q_2 , maka :

$$Q_2 = 60,5 + 13 \left(\frac{2,60 - 19}{12} \right) = 72,42$$

Untuk Q_3 , maka :

$$Q_3 = 73,5 + 13 \left(\frac{3,60 - 31}{23} \right) = 81,41$$

g. Desil

Desil : nilai pembagi 10 sama besar = D1, D2 D9

Jika sekelompok data dibagi menjadi 10 bagian yang sama banyak akan terdapat 9 pembagi, masing-masing disebut nilai desil (D) , D1, D2, D3,D9.

Nilai Desil ditentukan dengan rumus sebagai berikut:

Untuk Data tidak berkelompok:

$$D1 = \text{nilai yang ke-} \frac{i(n+1)}{10}, i = 1,2,3,,\dots,9$$

Untuk Data berkelompok

$$D1 = L_o + c \left\{ \frac{In/10 - F}{f} \right\}, i = 1,2,3,,\dots,9$$

Dimana

L_o : batas bawah kelas desil

c : lebar kelas

F : jumlah frekuensi semua kelas sebelum kelas desil

f : frekuensi kelas desil

Contoh soal

Data yang telah diurutkan

30, 35, 40, 45, 50, 55, 60, 65, 70, 80, 85, 95, 100

$$D1 = \text{nilai yang ke-} \frac{i(n+1)}{10}, i = 1,2,3,\dots,9$$

Maka D3 adalah

$$D3 = \text{nilai ke-} \frac{3(13+1)}{10} = 42/10 = \text{nilai ke } 4 \frac{1}{5}$$

Nilai ke $4 + \frac{1}{5}$ (nilai 5-nilai 4) = $45 + \frac{1}{5}(50-45) = 45 + 1 = 46$

Jadi D3 adalah 46

h. Persentil

Persentil : nilai pembagi 100 sama besar = P1, P2 P99

Jika sekelompok data dibagi menjadi 100 bagian yang sama banyak, maka akan terdapat 99 pembagi yang masing-masing disebut persentil P1,P2,P3,.....P99

Untuk Data tidak berkelompok:

$$P1 = \text{nilai yang ke-} \frac{i(n+1)}{100}, i = 1,2,3,\dots,99$$

Untuk Data berkelompok

$$P1 = L_o + c \left\{ \frac{In/100 - F}{f} \right\}, i = 1,2,3,\dots,9$$

Dimana

L_o : batas bawah kelas desil

c : lebar kelas

F : jumlah frekuensi semua kelas sebelum kelas persentil

f : frekuensi kelas persentil

4.3. PENUTUP

Latihan Soal

1. Berikut ini hasil nilai mata kuliah Statistik Deskriptif dan Inferensial mahasiswa fakultas Kesehatan :

85	95	73	60	90	70	82	60	95	40
66	50	70	65	90	75	88	66	100	45
70	40	66	66	90	70	78	65	100	50
80	60	83	68	90	75	88	70	65	10

Tentukan nilai :

- Mean, median, modus.
 - Kuartil K₁, K₂ dan K₃
2. Berikut ini data umur Ibu-ibu PKK Sayuk Rukun :

Umur (x)	fi
28-31	7
32-35	11
36-39	12
40-43	10
44-47	6
48-51	4

Tentukan nilai :

Tentukan tendensi sentral serta K₃, D₆, serta P₃₅

DAFTAR PUSTAKA

- Ating Somantri, Sambas Ali Muhidin, 2006. Aplikasi Statistika dalam Penelitian, Penerbit Pustaka Setia Bandung.
- Boediono, Wayan Koster, 2008. Teori dan Aplikasi statistika & Probabilitas, Remaja Rosdakarya, Bandung.
- Grace E.C. Korompis, 2014. Biostatistik Untuk Keperawatan, Penerbit Buku Kedokteran EGC.
- Husaini usman dkk, 2003. Pengantar Statistika, Bumi Aksara, Jakarta.

5

POKOK BAHASAN UKURAN PENYEBARAN DATA

5.1. PENDAHULUAN

5.1.1. Deskripsi Singkat

Pada bab ukuran penyebaran data yang akan dibahas adalah tentang menerapkan rumus ukuran penyebaran data yaitu standar deviasi, simpangan baku, varian, range, minimal & maksimal, kemiringan, kurtosis, distribusi normal.

5.1.2. Relevansi

Materi yang disajikan pada bab ini meliputi: standar deviasi, simpangan baku, varian, range, minimal & maksimal, kemiringan, kurtosis, distribusi normal.

5.1.3. Kompetensi

a. Standar Kompetensi

Pada akhir kuliah mahasiswa diharapkan dapat menjelaskan dan menerapkan rumus ukuran penyebaran data.

b. Kompetensi Dasar

Pada akhir kuliah diharapkan mahasiswa mampu:

1. Menggunakan rumus ukuran penyebaran data.
2. Menerapkan dalam soal kasus dengan menggunakan ukuran penyebaran data

5.2. PENYAJIAN

5.2.1. Uraian Isi

Ukuran penyebaran suatu kelompok data terhadap pusat data disebut **dispersi** atau **variasi** atau **keragaman data**. Adapun pentingnya mempelajari materi ini, yang pertama adalah pusat data seperti mean, median, modus hanya memberikan informasi yang sangat terbatas sehingga tanpa disandingkan dengan disperse data kurang bermanfaat dalam analisis data. Kedua, disperse data sangat penting untuk membandingkan penyebaran dua distribusi data atau lebih. Ukuran penyebaran ini menunjukkan suatu variasi dari suatu distribusi data. Dengan mengetahui variasi suatu data maka kita bias mengambil kesimpulan secara lebih tepat tentang distribusi Suatu data.

Rentang / Jangkauan (RANGE)

Nilai rentang ini menunjukkan selisih antara data yang paling tinggi dengan data yang paling rendah. Dengan melihat ukuran ini maka dapat diketahui gambaran secara kasar tentang variasi suatu distribusi data. Nilai range ini sangat kasar, karena tidak mempertimbangkan nilai-nilai yang lain selain nilai ekstrimnya. Adapun rumusnya adalah sebagai berikut :

$$\text{Range} = \text{max} - \text{min}$$

Contoh :

Pada data tabel satu diperoleh range pada masing-masing mata kuliah adalah :

Mata kuliah	Max – min	Range
A	50-50	0
B	75-25	50
C	85-10	70

Maka secara kasar dapat disimpulkan bahwa sebaran nilai pada mata kuliah C paling bervariasi dibandingkan pada mata kuliah A dan B. Nilai range sama dengan 0 pada mata kuliah A menunjukkan bahwa distribusi nilai A adalah homogen. Semakin besar nilai rentang maka distribusi data semakin bervariasi.

Rentang Antar Kuartil

Ukuran penyebaran yang dihitung dari selisih kuartil ketiga dan kuartil pertama merupakan variasi 50% populasi dengan menghilangkan 25% disisi kiri & 25% disisi kanan.

Dirumuskan :

$$\text{RAK} : K_3 - K_1$$

Simpangan Kuartil

Ukuran penyebaran yang dihitung dari setengah simpangan kuartil. Merupakan ukuran yang baik dibandingkan rentang antar kuartil.

Dirumuskan :

$$\text{SK} : \frac{1}{2} (K_3 - K_1)$$

Simpangan Rata-rata

Untuk menutup kekurangan dari nilai range maka bisa dihitung nilai simpangan rata-rata (mean deviation). Simpangan rata-rata (SR) memperhitungkan nilai-nilai lain selain nilai ekstrim distribusi data. SR merupakan jumlah nilai mutlak dari selisih semua nilai dengan nilai rata-rata dibagi banyaknya data. Adapun rumusnya adalah sebagai berikut :

Untuk data tidak berkelompok:

$$SR = \frac{\sum |X - \bar{X}|}{n}$$

dimana

X : nilai data

\bar{X} : rata-rata hitung

N : banyak data

untuk data berkelompok

$$SR = \frac{\sum f |X - \bar{X}|}{n} \quad \text{dimana : } n = \sum f$$

Tanda nilai mutlak pada rumus tersebut untuk menjamin agar simpangan bertanda positif, karena dispersi data merupakan ukuran yang positif.

Contoh soal:

Tentukan simpangan rata-rata kelompok data 20, 30, 50, 70, 80

Jawab

Rata-rata hitung $\bar{X} = 50$ dan $n = 5$ maka

$$SR = \frac{|20-50| + |30-50| + |50-50| + |70-50| + |80-50|}{5}$$

$$= \frac{30 + 20 + 0 + 20 + 30}{5} = \frac{100}{5} = 20$$

Jadi simpangan rata-rata adalah = 20

Variansi (Variance)

Simpangan baku atau *standard deviation* merupakan bentuk akar pangkat 2 dari variansi. Biasanya ukuran variansi ini diberi simpul sebagai S^2 (s pangkat 2). Sebenarnya yang merupakan ukuran simpangan adalah simpangan baku, namun demikian ukuran variansi ini merupakan ukuran pangkat dua dari simpangan baku, sehingga bisa juga dianggap sebagai ukuran penyebaran.

Variansi adalah rata-rata kuadrat selisih atau kuadrat simpangan dari semua nilai data terhadap rata-rata hitung. Variansi untuk sampel dilambangkan dengan S^2 , untuk populasi dilambangkan σ^2

Data tidak berkelompok :

$$S^2 = \frac{\sum (X - \bar{X})^2}{n - 1}$$

Data berkelompok :

$$S^2 = \frac{\sum f (X - \bar{X})^2}{n - 1}$$

Tentukan variansi dari kelompok data pada contoh soal simpangan rata-rata:

$$\begin{aligned} S^2 &= \frac{\sum (X - \bar{X})^2}{n - 1} \\ &= \frac{(20-50)^2 + (30-50)^2 + (50-50)^2 + (70-50)^2 + (80-50)^2}{5-1} \end{aligned}$$

$$\begin{aligned} & 900 + 400 + 0 + 400 + 900 \\ & = \frac{\text{-----}}{4} = 650 \end{aligned}$$

Jadi, variansi kelompok data tersebut adalah $S^2 = 650$

Simpangan Baku (Standard Deviation)

Simpangan baku ini merupakan ukuran penyebaran yang paling banyak digunakan. Ukuran ini dikenalkan oleh Karl Pearson. Dengan menggunakan simpangan rata-rata hasil pengamatan penyebaran sudah memperhitungkan seluruh nilai yang ada pada data. Namun demikian karena dalam penghitungan menggunakan nilai absolute maka tidak dapat diketahui arah penyebarannya. Maka dengan simpangan baku kelemahan ini dapat diatasi, yakni dengan cara membuat nilai pangkat 2, sehingga nilai negatif menjadi positif. Simpangan baku ini merupakan ukuran penyebaran yang paling teliti.

Standar deviasi berkaitan langsung dengan variansi. Standar deviasi adalah akar pangkat dua dari variansi.

Simpangan baku / standar deviasi :

Data tidak berkelompok :

$$S^2 = \sqrt{\frac{\sum (X - \bar{X})^2}{n - 1}}$$

Data berkelompok :

$$S^2 = \sqrt{\frac{\sum (f (X - \bar{X})^2)}{n - 1}}$$

Contoh :

Data tidak berkelompok dengan data sebagai berikut : 20, 30, 50, 70, 80

Maka diperoleh:

$$S^2 = 650$$

Maka standar deviasi adalah $\sqrt{650} = 25,495$

Atau menggunakan rumus berikut:

X	20	30	50	70	80	250
X ²	400	900	2500	4900	6400	15.100

$$\begin{aligned} \text{Variansi: } S^2 &= \sqrt{\frac{n \sum X - (\sum X)^2}{n(n-1)}} = \sqrt{\frac{5(15.100 - 250^2)}{5(4)}} \\ &= \sqrt{\frac{13000}{20}} \\ &= \sqrt{650} \end{aligned}$$

Standar deviasi $S = \sqrt{650} = 25,495$

Hasilnya akan sama antara contoh diatas dan dibawah.

MOMEN

Momen adalah penyimpanan data yang ke sekian, semisal momen ke-1, berarti penyimpangan data yang pertama rata-rata, momen ke-3 berarti penyimpangan data yang ke-3 dari rata-rata

Untuk data tidak berkelompok rumusnya :

$$M_r = \frac{\sum (X_i - \bar{X})^r}{n}$$

Dimana nilai r adalah nilai momen yang ke... (r= 1, 2, 3)

Untuk data berkelompok, rumusnya :

$$\text{Momen_ke-r} = \frac{\sum f_i \cdot X_i^r}{n}$$

Atau

$$M_r = \frac{\sum f_i (X_i - \bar{X})^r}{n}$$

kemudian ditentukan nilai momen sebagai berikut:

$$m_2 = m_2' - (m_1')^2$$

$$m_3 = m_3' - 3m_1'm_2' + 2(m_1')^3$$

$$m_4 = m_4' - 4m_1'm_3' + 6(m_1')^2 m_2' - 3(m_1')^4$$

KEMIRINGAN / SKEWNESS

Skewness: ukuran ketidaksimetrisan (kemencengan) distribusi. Distribusi yang ekor kurvanya lebih panjang kekanan disebut menceng ke kanan atau *positif skewness*. Begitu juga sebaliknya

Menggambarkan model kurva di distribusi data

- Model positif – ekor memanjang ke kanan
- Model negatif – ekor memanjang ke kiri
- Simetri – sisi kanan & kiri sama

Mo : Me : M

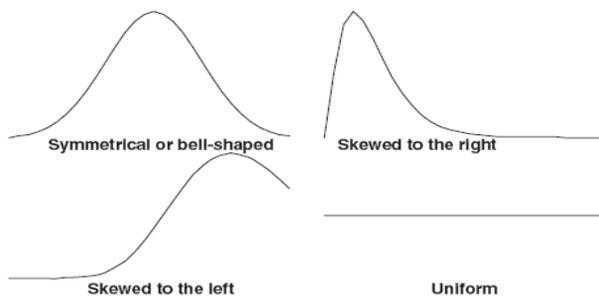


Fig. 2-3 Four common distributions.

1. Metode Karl Pearson

$$\text{Koefisien Skewness : SK : } \frac{\bar{X} - Mo}{S}$$

Karena $\bar{X} - Mo : 3 (X - Me)$

$$SK = \frac{3 (\bar{X} - Me)}{3}$$

Ket :

SK < 0 – Mdl negative

SK = 0 – simetri

SK > 0 – Mdl positif

Koef. Pearson I

$$\text{Skewness} = \frac{\text{mean} - \text{mode}}{\text{standard deviation}} = \frac{\bar{X} - \text{mode}}{s}$$

Koef. Pearson II

$$\text{Skewness} = \frac{3(\text{mean} - \text{median})}{\text{standard deviation}} = \frac{3(\bar{X} - \text{median})}{s}$$

Diperhatikan bila distribusinya normal maka koefisien skewness bernilai nol.

Koefisien skewness lainnya:

- *koef. kuartil skewness:*

$$\frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

- *koef. skewness 10-90% percentile:*

$$\frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{P_{90} - P_{10}} = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}}$$

- *koef. moment skewness:*

$$a_3 = \frac{m_3}{s^3} = \frac{m_3}{(\sqrt{m_2})^3} = \frac{m_3}{\sqrt{m_2^3}}$$

2. Metode Bowley

$$SK = \frac{(k_3 - k_2) - (k_2 - k_1)}{(k_3 - k_2) + (k_2 - k_1)}$$

- Simetri $k_3 - k_2 = k_2 - k_1$
Positif : $k_3 - k_2 > k_2 - k_1$
Negatif : $k_3 - k_2 < k_2 - k_1$

3. Nilai Skewness relatif

$$\beta_1 = (m_3^2 / m_2^3)$$

KERUNCINGAN DISTRIBUSI DATA

Keruncingan distribusi data adalah derajat atau ukuran tinggi rendahnya puncak suatu distribusi data terhadap distribusi normalnya data. Keruncingan distribusi data disebut juga kurtosis

Dikenal 3 bentuk :

- Lepto kurtik \rightarrow distribusi data yang puncaknya relatif tinggi
- Plati kurtik \rightarrow distribusi data yang puncaknya terlalu rendah atau terlalu mendatar.
- Meso kurtik \rightarrow distribusi data yang puncaknya tidak terlalu rendah dan tidak terlalu tinggi.

Macam-macam ukuran kurtosis:

- koef. moment kurtosis

$$a_4 = \frac{m_4}{s^4} = \frac{m_4}{m_2^2}$$

- ✓ kurtosis thd kuartil dan percentile

$$\kappa = \frac{Q}{P_{90} - P_{10}}$$

- kurtosis positif → distribusi lancip
- kurtosis negatif → distribusi tumpul

Dengan rumus momen

$\alpha_4 = 3$ maka keruncingan distribusi data disebut mesokurtis

$\alpha_4 > 3$ maka keruncingan distribusi data disebut leptokurtis

$\alpha_4 < 3$ maka keruncingan distribusi data disebut platikurtis

DISTRIBUSI NORMAL

Distribusi normal merupakan distribusi kontinu yang sangat penting dalam statistik dan banyak dipakai dalam memecahkan persoalan. Distribusi normal disebut **juga** distribusi Gauss.

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

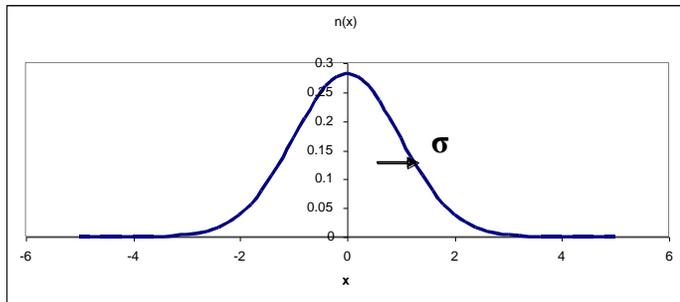
Dengan :

π = nilai konstanta, $\pi = 3,1416$

e = bilangan konstanta, $e = 2,7183$

μ = rata-rata populasi

σ = simpangan baku untuk distribusi



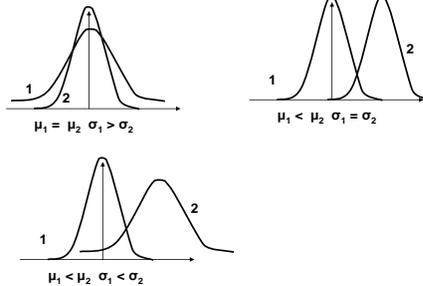
μ

sifat :

- 1) Grafik selalu ada di atas sumbu datar μ
- 2) Bentuknya simetrik terhadap $\mu = 0$

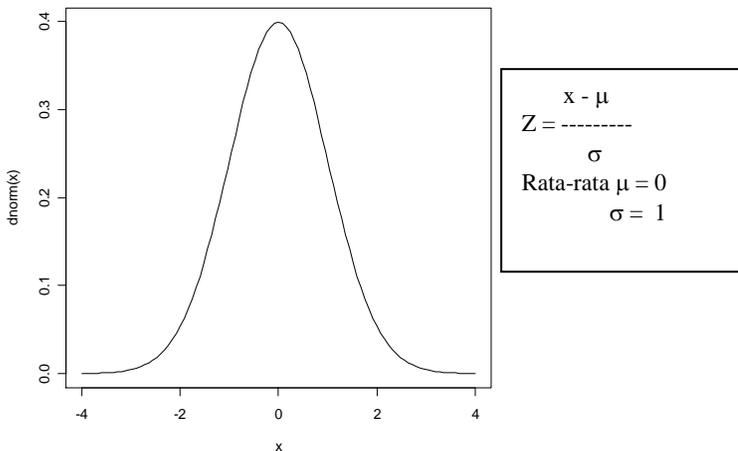
- 3) Mempunyai modulus (Unimodel)
- 4) Grafiknya mendekati sumbu μ dimulai dari $x = \mu + 3\sigma$ ke kanan dan $x = \mu - 3\sigma$ ke kiri
- 5) Luas daerah grafik selalu sama dengan satu unit persegi

Distribusi Normal : Sifat
Bentuk distribusi normal ditentukan oleh μ dan σ .



Dari distribusi **normal** di atas maka, dibuatlah suatu distribusi normal standar dengan mentransformasikannya.

Ditransformasi dengan :



Distribusi normal standar

Probabilitas \Leftrightarrow peluang

Diketahui suatu distribusi normal dengan $\mu = 50$ dan $\sigma = 10$

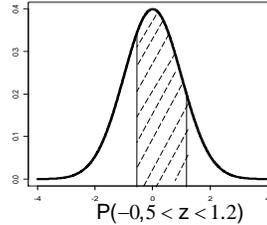
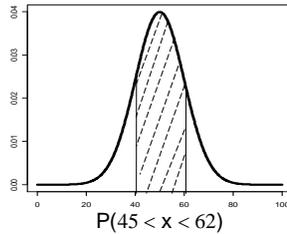
Carilah probabilitas bahawa X mendapat ilai antara 45 dan 62

Jawab:

Dicari nilai z yang berpadaan dengan $x_1 = 45$ dan $x_2 = 62$ adalah

$$z_1 = \frac{45-50}{10} = -0.5 \quad \text{dan} \quad z_2 = \frac{62-50}{10} = 1.2$$

Jadi: $P(45 < x < 62) = P(-0,5 < z < 1.2)$



Gunakan tabel distribusi normal standart, diperoleh:

$$P(45 < x < 62) = P(-0,5 < z < 1,2)$$

$$= P(z < 1,2) - P(z < -0,5)$$

$$= 0,8849 - 0,3085$$

$$= 0,5764$$

Dengan R

`> pnorm(-0.5)`

`[1] 0.3085375`

`> pnorm(1.2)`

`[1] 0.8849303`

Tabel 6.1. Luas daerah di bawah kurva normal

z	0.00	0.04	0.09
:					
:					
-0.5	0.3085				
0					
:					
:					
1.2	0.8849				
:					
:					

Contoh soal:

Variabel \bar{x} terdistribusi normal dengan mean 50 dan standard deviasi =10. Carilah probabilitas untuk menemukan \bar{x} bernilai antara 45 dan 62?

Jawab.

salam soal ini $\mu = 50$ dan $\sigma=10$. $x_1 = 45$ dan $x_2 =62$

Pertama kita mapping x ke z (melakukan normalisasi atau standardisasi):

$$z_1 = (x_1 - \mu) / \sigma \rightarrow z_1 = (45-50)/10 = -0.5$$

$$z_2 = (x_2 - \mu) / \sigma \rightarrow z_2 = (62-50)/10 = 1.2$$

Sehingga

$$P(45 < x < 62) = P(-0.5 < z < 1.2)$$

$$P(-0.5 < z < 1.2) = P(z < 1.2) - P(z < -0.5) = 0.8849 - 0.3085 = 0.5764$$

5.3. PENUTUP

Latihan Soal

1. Jelaskan yang dimaksud dengan simpangan rata-rata, variasi serta standar deviasi!
2. Diketahui terdapat 2 kelompok belajar dengan hasil sebagai berikut :
Kelompok data 1 : 7, 4, 10, 9, 15, 12, 12, 7, 9, 7, 8, 10, 13, 14, 15
Kelompok data 2 : 8, 11, 4, 3, 2, 5, 10, 6, 4, 1, 10, 8, 12, 6, 5.
Tentukan simpangan rata-rata, variasi dan standar deviasinya untuk masing-masing kelompok!
3. Apa yang dimaksud dengan derajat kemiringan distribusi data?
4. Sebutkan jenis-jenis kemiringan distribusi data!
5. Apa yang dimaksud dengan kurtosis?
6. Jelaskan jenis-jenis derajat keruncingan data?
7. Diketahui variabel X mempunyai distribusi normal dengan rata-rata 18 dan standar deviasi 2,5. Hitunglah:

- a. $P(X < 15)$
- b. $P(17 < X < 21)$
- c. Nilai k sehingga $P(X < k) = 0,2578$

DAFTAR PUSTAKA

Sudjana, Metode Statistika, Tarsito, Bandung, 1996

Sugiyono, Statistika untuk penelitian, Alfabeta, Bandung, 2002

Husaini usman dkk, Pengantar Statistika, Bumi Aksara, Jakarta, 2003

Djawarto, Statistik social ekonomi, BPFE, Yogyakarta, 1993

6

POKOK BAHASAN PENDUGAAN PARAMETER

6.1. PENDAHULUAN

6.1.1. Deskripsi Singkat

Pendugaan Parameter akan membahas tentang populasi dan sampel, cara menduga yang baik, pendugaan titik, pendugaan interval, pendugaan parameter populasi dengan sampel besar dan kecil.

6.1.2. Relevansi

Materi yang disajikan pada bab ini meliputi : populasi dan sampel, cara menduga yang baik, pendugaan titik, pendugaan interval, pendugaan parameter populasi dengan sampel besar dan kecil.

6.1.3. Kompetensi

a. Standar Kompetensi

Pada akhir kuliah mahasiswa diharapkan dapat menjelaskan dan mengaplikasikan Pendugaan Parameter.

b. Kompetensi Dasar

Pada akhir kuliah diharapkan mahasiswa mampu:

1. Melakukan pendugaan data dengan baik.
2. Menghitung pendugaan titik dan interval
3. Menghitung pendugaan parameter untuk sampel kecil dan besar.

6.2. PENYAJIAN

6.2.1. Uraian Isi

Pada bagian ini, mempelajari bagaimana cara melakukan pendugaan parameter populasi berdasarkan statistik yang dihitung dari sampel. Cara pengambilan kesimpulan tentang parameter yang pertama kali akan dipelajari ialah sehubungan dengan cara-cara menaksir harga parameter. Harga parameter yang sebenarnya tetapi tak diketahui itu akan ditaksir berdasarkan statistik sampel yang diambil dari populasi yang bersangkutan.

Penduga yang baik adalah merupakan penduga tidak bias, harapan penduga sama dengan yang diduga. Disamping itu penduga yang efisien, bila ada lebih dari satu penduga maka yang mempunyai variasi paling kecil. Penduga yang konsisten bila sampel diambil makin besar maka akan mendekati sesungguhnya. Adapun simbol θ adalah penduga populasi.

PARAMETER (POPULASI)	STATISTIK (SAMPEL)
Rerata Populasi (μ_x)	Rerata Sampel (X)
Varian Populasi (σ_x^2)	Varian Sampel (S_x^2)
Standar Deviasi Pop. (σ_x)	Standar Deviasi Sampel (S_x)

PENDUGAAN TITIK

Dalam menentukan pendugaan ada dua jenis pendugaan yaitu pendugaan titik dan pendugaan interval. Bila nilai parameter θ dari populasi hanya diduga dengan memakai satu nilai statistik dari sampel yang diambil dari populasi maka disebut pendugaan titik. Semakin dekat nilai penduga dengan nilai yang diduga maka penduga akan semakin baik. Misalkan bila diambil sampel 100 mahasiswa didapatkan bahwa rata-rata berat badan mahasiswa fakultas kesehatan

adalah 45 kg sedangkan hasil pengambilan rata-rata populasi mahasiswa Universitas Dian Nuswantoro 46 kg. Sehingga ada dua nilai populasi dan sampel yang berbeda, ada keraguan dalam menentukan rata berat badan mahasiswa yang sesungguhnya. Pendugaan titik memiliki kelemahan dan sulit dipertanggungjawabkan secara statistik.

PENDUGAAN INTERVAL

Bila nilai parameter θ dari populasi diduga dengan memakai beberapa nilai statistik θ yang berada dalam suatu interval, sehingga dapat digambarkan sebagai berikut : $\theta_1 < \theta < \theta_2$, maka statistik θ disebut penduga interval.

Contoh : rata-rata berat badan mahasiswa Fakultas Kesehatan diduga memakai interval $50 < \theta < 60$, artinya rata-rata berat badan mahasiswa terletak pada interval tersebut. Kita dapat menduga juga bahwa berat badan mahasiswa Fakultas Kesehatan antara $45 < \theta < 65$. Makin lebar interval, makin besar kepercayaan atau keyakinan kita bahwa rata-rata berat badan mahasiswa terletak pada interval tersebut. Dalam praktiknya kita sebaiknya memakai suatu interval yang sempit, tetapi mempunyai derajat kepercayaan atau derajat keyakinan yang dapat diterima.

Derajat penduga θ disebut koefisien kepercayaan yang ditulis dengan α dimana $0 < \alpha < 1$ dan dinyatakan dalam bentuk probabilitas. Misalnya $P(\theta_1 < \theta < \theta_2) = 0,95$ artinya dengan probabilitas 0,95 bahwa sampel acak yang kita ambil akan menghasilkan suatu interval $\theta_1 < \theta < \theta_2$ yang mengandung parameter θ dari populasi. Contohnya dalam berat badan $50 < \theta < 60 = 0,95$; dalam statistika biasanya yang dipilih adalah interval yang lebih pendek, tetapi dengan

probabilitas yang tinggi atau dengan derajat kepercayaan yang tinggi.

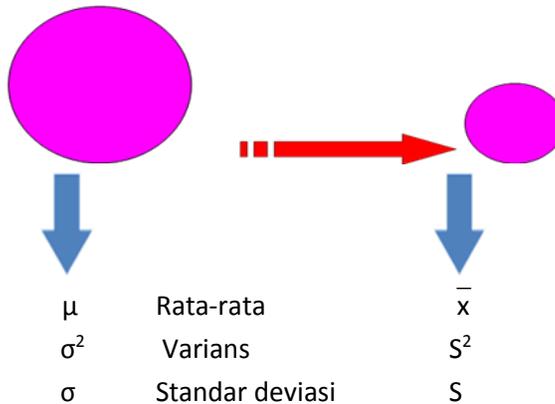
$$P(\bar{\theta}_1 < \theta < \theta_2) = 1 - \alpha, 0 < \alpha < 1$$

α : koefisiensi kepercayaan;

$(1 - \alpha)$ disebut derajat kepercayaan

PENDUGAAN PARAMETER POPULASI DENGAN SAMPEL BESAR

Bila pada suatu populasi diambil sampel acak yang besar, maka statistik θ akan mempunyai distribusi normal, sehingga dapat ditransformasi menjadi distribusi normal standar.



Penaksiran dilakukan diantara 2 nilai estimasi, ada batas bawah & batas atas berdasarkan interval kepercayaan tertentu. Semakin tinggi interval kepercayaan yg digunakan, maka interval semakin baik

Interval kepercayaan 90% \rightarrow 1,64

Interval kepercayaan 95% \rightarrow 1,96

Interval kepercayaan 99% \rightarrow 2,58

Semakin sempit interval yang dihasilkan dalam estimasi, maka penaksiran presisi (semakin tepat).

a. Pendugaan Parameter μ

$\sigma_{\bar{x}} = (\mathbf{x}/(\sqrt{n}))$, bila populasi tak terbatas

$\sigma_{\bar{x}} = (\mathbf{x}/(\sqrt{n}) \sqrt{((N - n)/(N - 1))})$, bila populasi terbatas

Contoh :

Dari populasi para pegawai suatu perusahaan diambil sampel sebanyak 100 orang dan dicatat gaji tahunan masing-masing. Rata-rata dan simpangan baku gaji mereka itu adalah: $\bar{x} = 30.000.000$ $S = 6.000.000$. buatlah selang kepercayaan 95% untuk menduga berapa sesungguhnya rata-rata gaji para pegawai di perusahaan tersebut!

Jawab :

Populasi dianggap tak terbatas, sebab ukurannya tidak diketahui.

Sampel : $n = 100$, $\bar{x} = 30.000.000$ dan $S = 6.000.000,-$. Ukuran sampel $n = 100$ cukup besar. Karena σ tidak diketahui, maka harus ditaktir dengan S , yaitu:

$$\sigma_{\bar{x}} = (\mathbf{x}/(\sqrt{n})) = \frac{6.000.000}{\sqrt{100}} = 600.000$$

Untuk internal kepercayaan 95%, diperoleh $Z_{\alpha/2} = 1,96$

Maka :

$$\begin{aligned} & \bar{X} \pm Z_{1/2\gamma} \times \left[\frac{\sigma}{\sqrt{n}} \right] \\ & = 30.000.000 + (1,96) \times 600.000 = 31.176.000 \\ & = 30.000.000 - (1,96) \times 600.000 = 28.824.000 \end{aligned}$$

Jadi, interval kepercayaan 95% untuk rata-rata (μ) gaji tahunan yang sesungguhnya dari para pegawai di perusahaan tersebut adalah $P 28.824.000 < \mu < 31.176.000 = 0,95$

Artinya, kita percaya 95% bahwa rata-rata gaji tahunan yang sesungguhnya dari para pegawai di perusahaan itu berkisar antara Rp 28.824.000,- sampai dengan Rp 31.176.000,- setahun.

b. Pendugaan Parameter Proporsi

Bila suatu populasi berukuran N mengandung jenis tertentu dengan proporsi $p = \frac{X}{N}$ dan populasi itu diambil secara berulang sampel berukuran n yang mengandung jenis tertentu dengan $\bar{p} = \frac{X}{n}$, maka distribusi sampel proporsi \bar{p} akan mempunyai rata-rata $\mu_p = p$ dan simpangan baku:

$$\pi = p \pm Z_{1/2\gamma} \times \sqrt{\left[p \left(\frac{1-p}{n} \right) \right]}$$

Contoh:

Suatu penelitian ingin menaksir persentase absensi di karyawan yang berjumlah 1500 orang diambil sejumlah 70 karyawan, ternyata tingkat absensinya mencapai 10%. Taksirlah tingkat absensi pada derajat 90%.

$$\begin{aligned} \pi &= p \pm Z_{1/2\gamma} \times \sqrt{\frac{p(1-p)}{n}} \\ &= 0,1 \pm 1,64 \times \sqrt{\frac{0,1(0,9)}{70}} \end{aligned}$$

$$= 0,1 \pm 0,06$$

$$0,04 < \mu < 0,16$$

Artinya kita percaya 90% bahwa proporsi absensi karyawan antara $0,04 < \mu < 0,16$.

c. Pendugaan Parameter Beda Dua Rata-rata ($\mu_1 - \mu_2$)

Populasi pertama mempunyai rata-rata μ_1 dan simpangan baku σ_1 ; sedangkan populasi kedua mempunyai rata-rata μ_2 dan simpangan baku σ_2 . Dari populasi pertama kita ambil sampel acak sebanyak n_1 dan dari populasi kedua n_2 , kemudian kita hitung rata-rata \bar{x}_1 untuk sampel pertama

dan rata-rata \bar{x}_2 untuk sampel kedua. Misalkan kedua sampel itu saling bebas. Bila kedua sampel acak itu diambil secara berulang, maka kita akan memperoleh distribusi sampel beda dua rata-rata ($\bar{x}_1 - \bar{x}_2$) dengan rata-rata ($\mu_1 - \mu_2$)

dan simpangan baku $\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Contoh:

Ujian kalkulus diberikan kepada dua kelompok mahasiswa, yaitu mahasiswa perempuan sebanyak 75 orang dan mahasiswa laki-laki sebanyak 50 orang. Kelompok mahasiswa perempuan memperoleh nilai rata-rata 82 dengan simpangan baku 8, sedangkan kelompok mahasiswa laki-laki memperoleh nilai rata-rata 76 dengan simpangan baku 6. Bila μ_1 menyatakan rata-rata nilai ujian kelompok mahasiswa perempuan dan μ_2 menyatakan rata-rata nilai ujian kelompok mahasiswa laki-laki, buatlah interval kepercayaan 96% untuk menduga berapa sesungguhnya beda rata-rata dua kelompok mahasiswa tersebut!

Jawab:

Dua populasi dianggap tak terbatas

Kelompok mahasiswa perempuan : $n_1 = 75$ $\bar{x}_1 = 82$, $S_1 = 8$

Kelompok mahasiswa laki-laki : $n_2 = 50$, $\bar{x}_2 = 76$, $S_2 = 6$

Dalam hal ini simpangan baku dua populasi mahasiswa itu tidak diketahui, maka simpangan sampel dua rata-rata tersebut adalah:

$$\begin{aligned} \sigma_{\bar{x}_1 - \bar{x}_2} &= \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \\ &= \sqrt{\frac{8^2}{75} + \frac{6^2}{50}} \\ &= 1,254 \end{aligned}$$

Penduga untuk $(\mu_1 - \mu_2)$ adalah $(\bar{x}_1 - \bar{x}_2)$

Untuk interval kepercayaan 96%, maka $Z_{\alpha/2} = 2,05$ sehingga diperoleh:

$$(\bar{x}_1 - \bar{x}_2) - Z_{\alpha/2} \sigma_{\bar{x}_1 - \bar{x}_2} = (82 - 76) - (2,05) \times (1,254) = 3,429$$

$$(\bar{x}_1 - \bar{x}_2) + Z_{\alpha/2} \sigma_{\bar{x}_1 - \bar{x}_2} = (82 - 76) + (2,05) \times (1,254) = 8,571$$

Jadi, interval kepercayaan 96% untuk penduga $(\mu_1 - \mu_2)$ adalah $P(3,429 < \mu_1 - \mu_2 < 8,571) = 0,96$

Artinya 96% dapat dipercaya bahwa beda sesungguhnya nilai rata-rata ujian kalkulus dua kelompok mahasiswa itu terletak antara 3,429-8,571.

d. Pendugaan Parameter Beda Dua Proporsi ($p_1 - p_2$)

Populasi pertama mengandung jenis tertentu dengan

proporsi $P_1 = \frac{X_1}{N_1}$ dan populasi kedua mengandung jenis tertentu dengan proporsi $p_2 = \frac{X_2}{N_2}$. Bila pada dua populasi

diambil sampel acak masing-masing n_1 dan n_2 , maka sampel pertama akan mengandung jenis tertentu dengan proporsi

$p_1 = \frac{X_1}{n_1}$ dan sampel kedua akan mengandung jenis tertentu

$p_2 = \frac{X_2}{n_2}$. Bila sampel diambil secara berulang dan

saling bebas, maka akan diperoleh distribusi sampel beda dua proporsi ($p_1 - p_2$). Dengan demikian interval kepercayaan untuk penduga beda dua proporsi ($p_1 - p_2$):

PENDUGAAN PARAMETER POPULASI DENGAN SAMPEL KECIL

Pendugaan itu berlaku untuk populasi berdistribusi normal maupun tidak normal. Jarang sekali variansi σ^2 dari suatu populasi diketahui. Akan tetapi, bila sampel yang kita ambil

bersifat acak dan berukuran besar maka σ^2 dapat ditaksir dengan variansi yang dihitung dari sampel, yaitu:

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

Dengan sampel yang besar, maka fluktuasi S^2 tidak akan terlalu besar, artinya nilai-nilai S^2 tidak akan terlalu berbeda antara sampel yang satu dengan sampel yang lain. Sehingga variasi σ^2 dari populasi dapat didekati dengan variasi dari sampel, yaitu S^2 , karena S^2 merupakan penduga yang baik untuk σ^2 . Dalam hal ini, apapun distribusi populasinya, normal atau tidak normal, maka statistik:

$$\frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

dimana $\sigma / \sqrt{n} = (s / (\sqrt{n}))$

Dalam hal sampel yang kita ambil jumlahnya kecil, ternyata distribusi dari statistik tersebut merupakan distribusi student yang ditulis t yaitu :

$$t = \frac{(\bar{x} - \mu) / (s / \sqrt{n})}{(df)}$$

6.3. PENUTUP

Latihan

1. Seorang Kepala Puskesmas ingin mengetahui lebih lanjut persentase Ibu yang mengimunitasikan anaknya dengan tertib pada tahun 2015. Dari sampel yang dikumpulkan sebanyak 200 Ibu, ternyata ada 15 ibu yang tidak lengkap mengimunitasikan anaknya. Buatlah interval kepercayaan 90% untuk menduga berapa sesungguhnya Ibu yang tidak lengkap mengimunitasikan anaknya!
2. Dalam suatu sampel acak yang terdiri atas 40 laki-laki dan 60 perempuan yang mengikuti seminar “Bahaya Merokok”

ternyata terdapat sebanyak 35 orang laki-laki dan 25 orang perempuan yang puas dan senang dengan acara seminar tersebut. Buatlah interval kepercayaan 90% untuk memperkirakan beda proporsi antara banyaknya laki-laki dan perempuan yang sesungguhnya menyukai acara seminar tersebut!

DAFTAR PUSTAKA

- Ating Somantri, Sambas Ali Muhidin, 2006. Aplikasi Statistika dalam Penelitian, Penerbit Pustaka Setia Bandung.
- Boediono, Wayan Koster, 2008. Teori dan Aplikasi statistika & Probabilitas, Remaja Rosdakarya, Bandung.
- Grace E.C. Korompis, 2014. Biostatistik Untuk Keperawatan, Penerbit Buku Kedokteran EGC.

7

POKOK BAHASAN HIPOTESIS

7.1. PENDAHULUAN

7.1.1. Deskripsi Singkat

Pada bab Hipotesis akan membahas tentang pengertian hipotesis dan prinsip pengujiannya, bentuk uji hipotesis, identifikasi kesalahan pengambilan keputusan, tingkat kemaknaan dalam pengujian hipotesis serta jenis-jenis pengujian hipotesis.

7.1.2. Relevansi

Materi yang disajikan pada bab ini meliputi : pengertian hipotesis dan prinsip pengujiannya, bentuk uji hipotesis, identifikasi kesalahan pengambilan keputusan, tingkat kemaknaan dalam pengujian hipotesis serta jenis-jenis pengujian hipotesis

7.1.3. Kompetensi

a. Standar Kompetensi

Pada akhir kuliah mahasiswa diharapkan dapat menjelaskan dan menentukan Hipotesis.

b. Kompetensi Dasar

Pada akhir kuliah diharapkan mahasiswa mampu:

1. Menjelaskan pengertian hipotesis dan prinsip pengujiannya
2. Membedakan arah dan bentuk uji hipotesis
3. Mengidentifikasi kesalahan pengambilan keputusan dalam pengujian hipotesis
4. Menentukan tingkat kemaknaan dalam pengujian hipotesis dan menguraikan jenis pengujian hipotesis

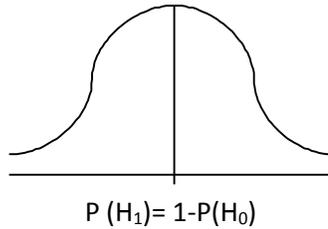
7.2. PENYAJIAN

7.2.1. Uraian Isi

Pengujian hipotesis membantu pengambilan keputusan dalam menolak atau menerima suatu hipotesis yang diajukan. Keyakinan menolak atau menerima berdasarkan pada besarnya peluang untuk memperoleh hubungan secara kebetulan. Prinsip uji hipotesis adalah melakukan perbandingan nilai sampel atau data hasil penelitian dengan nilai hipotesis atau nilai populasi yang diajukan.

PENGERTIAN HIPOTESIS

HIPOTESIS (HYPOTHESIS) Berasal dari bahasa Yunani, Hupo : sementara ;Thesis=pernyataan/dugaan. Karena merupakan pernyataan sementara maka hipotesis harus diuji kebenarannya. Hipotesis terbagi dua yaitu hipotesis penelitian (*reseach hypothesis*) dan hipotesis statistik (*Statistical hypithesis*). Kriteria menterjemahkan dugaan penelitian ke dalam hipotesis statistik dalam bentuk H_0 dan $H_1(H_a)$. H_0 dan $H_1(H_a)$ ini bersifat komplementer artinya apa yang ada dalam H_0 tidak terdapat dalam H_1 dan sebaliknya dalam notasi:



Hipotesis nol (H_0) menyatakan bahwa tidak ada perbedaan kejadian diantara dua kelompok atau tidak ada hubungan antara satu variabel dan variabel yang lain. Sedangkan alternatif (H_a) menyatakan bahwa terdapat perbedaan suatu kejadian diantara dua kelompok atau terdapat hubungan satu variabel dengan variabel yang lain.

ARAH DAN BENTUK UJI HIPOTESIS

Bentuk hipotesis alternatif dapat menentukan arah uji statistik yaitu satu arah (one tail) atau dua arah (two tail)

1. Satu arah dipilih bila hipotesis alternatif menyatakan bahwa terdapat perbedaan satu sama lain lebih tinggi atau lebih rendah. Contoh hipotesis penelitian:

- Pekerja perokok lebih cepat lelah dibandingkan yang tidak merokok.
- Ada dugaan bahwa secara rata-rata tingkat partisipasi masyarakat desa dalam pembangunan lebih tinggi dari pada rata-rata tingkat partisipasi masyarakat kota.

Contoh hipotesis statistik:

$H_0 : \mu_D \geq \mu_K$ dan $H_a \mu_D < \mu_K$, Perhatikan tanda lebih besar pada H_a tanda tersebut menunjukkan uji hipotesis satu arah, yaitu ke sebelah kiri

2. Dua arah merupakan hipotesis alternatif yang hanya menyatakan perbedaan tanpa melihat tinggi rendahnya perbedaan. Sebagai contoh hipotesis penelitian:
- Ada perbedaan tingkat kelelahan pada pekerja perokok dan tidak perokok.
 - Ada dugaan bahwa secara rata-rata tingkat partisipasi masyarakat desa dalam pembangunan berbeda dengan rata-rata tingkat partisipasi masyarakat.

Contoh hipotesis statistik :

Berdasarkan dugaan penelitian tsb kita bisa menterjemahkan dalam H_0 dan H_1 seperti berikut:

$$H_0 : \mu_D = \mu_K \text{ dan } H_1 \mu_D \neq \mu_K$$

Tanda tidak sama dengan, menunjukkan uji hipotesis berlangsung dua arah, yaitu sebelah kiri dan sebelah kanan yang artinya bahwa daerah dan titik kritis ada dibelah kiri dan sebelah kanan.

KESALAHAN PENGAMBILAN KEPUTUSAN

KEPUTUSAN PENGUJIAN	KEADAAN SEBENARNYA	
	HO BENAR	HO SALAH
MENOLAK H_0	KESALAHAN TIPE I (α)	KEPUTUSAN BENAR ($1-\beta$)
MENDUKUNG H_0	KEPUTUSAN BENAR ($1-\alpha$)	KESALAHAN TIPE II (β)

Terdapat kesalahan dalam pengambilan keputusan:

1. Kesalahan Tipe I (α)
 Kesalahan ini terjadi karena menolak H_0 , padahal sesungguhnya H_0 benar. Hal ini berarti menyimpulkan adanya perbedaan padahal sesungguhnya tidak ada beda. Peluang kesalahan tipe I atau α atau tingkat kemaknaan (significance level), sebaliknya peluang

untuk membuat kesalahan tipe I adalah sebesar $1-\alpha$ disebut tingkat kepercayaan (confidence level).

2. Kesalahan Tipe II (β)

Kesalahan ini terjadi karena tidak menolak H_0 , padahal sesungguhnya H_0 salah. Hal ini menyimpulkan tidak ada perbedaan, padahal sesungguhnya terdapat perbedaan. Peluang untuk membuat kesalahan tipe II ini sebesar β . Peluang untuk tidak membuat kesalahan tipe II adalah sebesar $1-\beta$, dan dikenal dengan tingkat kekuatan uji (power of the test).

MENENTUKAN TINGKAT KEMAKNAAN

Tingkat kemaknaan atau sering disebut nilai α , merupakan nilai yang menunjukkan besarnya peluang salah dalam menolak hipotesis (H_0). Nilai α merupakan nilai batas maksimal kesalahan menolak H_0 . Bila H_0 ditolak berarti ada perbedaan atau ada hubungan. Penentuan nilai α bergantung pada tujuan dan kondisi penelitian. Nilai α yang sering digunakan adalah 10%, 5% atau 1% untuk bidang kesehatan masyarakat biasanya digunakan nilai α sebesar 5%.

LANGKAH-LANGKAH DALAM PENGUJIAN HIPOTESIS

- Tetapkan dulu rumusan hipotesis dengan tepat, baik hipotesis nol (H_0) dan hipotesis alternatif (H_a) apakah termasuk uji satu arah atau uji dua arah.
- Tetapkan taraf nyata α yang diinginkan sehingga dapat diperoleh nilai kritis dalam tabel dengan demikian dapat digambarkan daerah penolakan atau penerimaan H_0
- Tetapkan statistik uji yang cocok untuk menguji hipotesis nol. Rumus statistik uji sangat tergantung pada parameter populasi yang diuji.

- Hitunglah nilai statistik uji berdasarkan data & informasi yang diketahui baik dari populasi maupun dari sampel yang diambil dari populasi.
- Simpulkan tolak H_0 bila nilai statistik uji jatuh atau terletak pada didaerah penolakan H_0 bilamana $Z_h > Z_\alpha$ atau
 $Z_h < -Z_\alpha$ untuk uji satu arah $Z_h > Z_{\alpha/2}$ atau
 $Z_h < -Z_{\alpha/2}$ untuk uji dua arah

JENIS UJI HIPOTESIS

Menguji Beda Mean Satu Sampel

Bila nilai σ diketahui, digunakan uji Z dengan rumus :

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Bila nilai σ tidak diketahui, digunakan dengan uji t dengan rumus:

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

Keterangan

- \bar{x} : mean data sampel
- μ : mean data populasi
- σ : standar deviasi data populasi
- S : standar deviasi data sampel
- n : jumlah sampel yang diteliti

contoh:

Data dari pihak akademik menyatakan bahwa rata-rata tinggi badan mahasiswa 155 cm untuk menguji pernyataan itu diambil 49 mhs & diukur TB 152 cm. Diketahui

simpangan baku dari populasi adalah 3,2 cm. Ujilah pada tingkat 5% apakah betul bahwa rata2 TB adalah 155 cm.

- Diketahui:

$$\mu = 155$$

$$\bar{x} = 152$$

$$n = 49$$

$$\sigma = 3,2$$

$$\alpha = 5\%$$

- Hipotesis

$$H_0 = \mu = 155$$

$$H_a = \mu \neq 155$$

- Titik kritis z tabel $\alpha = 5\%$ yaitu $\pm 1,96$

- Penentuan keputusan

H_0 diterima bila $-1,96 \leq Z \leq 1,96$

H_0 ditolak bila $z < -1,96$ atau $Z \geq 1,96$

- Z Hitung

$$Z \text{ hitung} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{152-155}{3,2 / \sqrt{49}} = \frac{-3}{0,457} = -6,56$$

- Keputusan

Karena jatuh didaerah penolakan H_0 , maka H_0 ditolak dan H_a diterima. Kesimpulan rata-rata tinggi badan mahasiswa tidak sama dengan 155 pada tingkat signifikansi 5 %.

UJI BEDA DUA RATA-RATA UNTUK SAMPEL KECIL ($n < 30$)

Rumusnya :

T hitung =

$$\frac{(x_1 - x_2) / \sqrt{((n_1 - 1)(SD_1^2) + (n_2 - 1)(SD_2^2)) / (n_1 + n_2 - 2)}}{\sqrt{(1/n_1 + 1/n_2)}}$$

Keterangan :

\bar{x}_1 : rata-rata statistik untuk sampel pertama

\bar{x}_2 : rata-rata statistik untuk sampel kedua

SD₁ : standar deviasi untuk sampel pertama

SD₁²: varian sampel 1

SD₂ : standar deviasi untuk sampel kedua

SD₂²: varian sampel 2

n₁ : jumlah sampel pertama

n₂ : jumlah sampel kedua

UJI BEDA DUA RATA-RATA UNTUK SAMPEL BESAR (n>30)

Rumus:

$$Z_{\text{hitung}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{SD_1^2}{n_1} + \frac{SD_2^2}{n_2}}}$$

UJI BEDA PROPORSI

$$Z_{\text{hitung}} = \frac{X - (n \cdot \pi)}{\sqrt{(n \cdot \pi \cdot q)}}$$

Keterangan

X : nilai sampel yang diketahui dari pengamatan

π : proporsi dari parameter

$$q = 1 - \pi$$

n : jumlah sampel yang digunakan.

Jika proporsi (P) dihitung dengan menggunakan rumus

X/n maka rumus tersebut dapat diubah menjadi

$$Z_{\text{hitung}} = \frac{(P - \pi)}{\sqrt{(\pi \cdot q)/n}}$$

7.3. PENUTUP

Latihan

1. Kepala Puskesmas Sehat di Kota Bahagia melaporkan bahwa rata-rata berat bayi saat lahir tahun lalu adalah 3 kg dengan standar deviasi 300 gram. Kepala Puskesmas Sehat ingin menguji apakah ada perbedaan rata-rata berat bayi tahun lalu dengan saat ini, untuk menguji hal tersebut diambil sampel sebanyak 100 bayi dan diperoleh rata-rata beratnya 3,2 kg. Buktikan apakah ada perbedaan rata-rata berat bayi tahun lalu dengan saat ini ($\alpha : 5\%$)
2. Berdasarkan hasil survei di kota Damai, dilaporkan bahwa 25% remaja merokok. Saat ini dilakukan survei pada 250 remaja, ternyata ditemukan 100 remaja merokok. Buktikan apakah persentase remaja merokok tahun ini lebih banyak dibandingkan tahun lalu!

8

POKOK BAHASAN VALIDITAS DAN RELIABILITAS

8.1. PENDAHULUAN

8.1.1. Deskripsi Singkat

Dalam pokok bahasan ini akan menjelaskan mengenai validitas dan reliabilitas. Pengertian validitas dan reliabilitas, macam validitas dan reliabilitas serta syarat pengujian validitas dan realibilitas instrumen penelitian.

8.1.2. Relevansi

Materi yang akan dibahas dalam pokok bahasan ini adalah pengertian, macam dan syarat menguji validitas dan reliabilitas instrumen penelitian.

8.1.3. Kompetensi

Pada akhir perkuliahan mahasiswa diharapkan mampu:

1. Menjelaskan pengertian validitas
2. Menjelaskan pengertian realibilitas
3. Membedakan validitas dan reliabilitas

8.2. PENJELASAN

8.2.1. Uraian Isi

A. Validitas

a. Pengertian Validitas

Validitas berasal dari kata *validity* yang memiliki arti ketepatan & kecermatan. Valid atau sah bila alat ukur itu benar-benar mengukur apa yang hendak diukur. Contoh meteran digunakan untuk mengukur panjang, timbangan digunakan untuk mengukur berat, literan digunakan untuk mengukur volume. Macam-macam validitas:

1. Menurut sugiyono, validitas terbagi 2:

a. Validitas Dalam

b. Validitas Luar

2. Pratiknya membagi 3

a. Validitas isi

b. Validitas konstruksi

c. Validitas kriterium

b. Validitas Isi

Tingkat representativitas isi/substansi pengukuran terhadap konsep variabel sebagaimana dirumuskan dalam definisi operasional. Artinya kalo alat ukur berupa pertanyaan (kuesioner) maka kalimat pertanyaan itu mewakili substansi apa yang hendak diukur. Contoh, Mengukur tingkat pengetahuan ibu mengenai gizi balita, bukan gizi orang dewasa

c. Validitas Kriterium

Contohnya Bila seorang mahasiswa ketika test masuk mendapat nilai tinggi kemudian selama mengikuti kuliah sampai dengan tamat, ternyata lancar & mudah dengan nilai-nilai yang bagus, maka alat ukur berupa test ujian

masuk PT memiliki validitas prediksi yang handal. Bila analisis test korelasi kuat berarti validitas kriterium tinggi dan sebaliknya.

d. Validitas Muka

Berkaitan dengan pengukuran atribut yang konkrit, tanpa membuat suatu inferensi atau suatu kesimpulan. Bila berkaitan dengan para ahli seorang peneliti membuat alat ukur untuk mengukur skala contohnya perilaku membuang sampah padat yang sehat kemudian dikonsulkan pada ahli perilaku sehat berkaitan dengan kesehatan lingkungan.

e. Validitas Internal

Alat ukur yang telah memiliki validitas konstruksi yang tinggi dan validitas isi yang tinggi. Bahwa untuk instrumen yang nontest digunakan untuk mengukur sikap cukup memenuhi validitas konstruksi

f. Validitas Eksternal

Bahwa bila kriteria dalam instrumen terdapat kesamaan dengan kriteria dilapangan atau fakta-fakta empiris, maka alat ukur itu memiliki validitas eksternal yang tinggi Contohnya untuk mengukur kinerja paramedis di RS tertentu, bila pada instrumen dibandingkan dengan catatan-catatan RS tentang kinerja paramedis maka bila terdapat kesamaan antara kriteria dalam instrumen dgn fakta di lapangan maka validitas eksternal yang tinggi.

g. Validitas Konstruk

Validitas konstruk ditekankan pada konstruksi pertanyaan satu dengan lainnya memiliki hubungan yang erat satu sama lainnya Pertanyaan satu dengan lainnya atau pokok-pokok yang dicantumkan dalam instrumen satu sama lainnya bergayut atau relevan.

Contoh: Pengetahuan tentang gizi balita ibu2 hamil yang datang ke RS maka pokok masalah yakni pengetahuan tentang gizi balita.

B. Syarat Uji Validitas

- a. Paling sedikit 30 responden, dengan ciri responden uji coba harus mirip ciri-cirinya dengan responden penelitian.
- b. Alasan 30 responden adalah batas jumlah antara sedikit & banyak, dengan pengertian bahwa data diatas 30 kurva akan mendekati kurva normal.
- c. Hasil uji coba dilakukan uji korelasi antara skor item dengan skor total. Bila korelasinya rendah berarti pertanyaan itu tidak bergayut & harus didrop

C. Reliabilitas

Dalam menguji reliabilitas dibagi menjadi tiga yaitu:

- a. Teknik test-retest
Instrumen diujikan pada responden yang sama, dalam selang waktu antara kira-kira antara 15-30 hari. Bila terlalu dekat kurang baik sebab masih ingat betul jawaban pertama, bila terlalu lama kurang bagus karena mungkin sudah terjadi perubahan pada diri responden dalam hal variabel yang hendak diukur. Hasil pengukuran pertama dikorelasikan dengan hasil pengukuran yang kedua menggunakan rumus product momen, bila signifikan berarti reliabel, bila tidak signifikan tidak reliabel.
- b. Teknik Belah Dua
Melakukan uji coba alat ukur pada sejumlah responden kemudian dihitung validitasnya. Item-item yang valid

dikumpulkan yang tidak valid dibuang. Item-item yang valid tersebut dibagi dua. Skor masing-masing item tiap belahan dijumlahkan, sehingga menghasilkan dua skor total dari belahan kedua dengan menggunakan teknik product moment. Oleh karena hasil korelasi berasal dari angka-angka item yang dibelah maka angka korelasinya akan lebih rendah daripada tidak dibelah. Oleh karena itu harus dicari angka korelasi untuk seluruh item. Dengan menggunakan rumus

c. Teknik Bentuk Paralel

Melakukan uji coba alat ukur pada sejumlah responden kemudian dihitung validitasnya. Item-item yang valid dikumpulkan yang tidak valid dibuang. Item-item yang valid tersebut dibagi dua. Skor masing-masing item tiap belahan dijumlahkan, sehingga menghasilkan dua skor total dari belahan kedua dengan menggunakan teknik product moment. Oleh karena hasil korelasi berasal dari angka-angka item yang dibelah maka angka korelasinya akan lebih rendah daripada tidak dibelah. Oleh karena itu harus dicari angka korelasi untuk seluruh item.

8.3. PENUTUP

Latihan

1. Jelaskan apa yang dimaksud dengan validitas?
2. Jelaskan jenis-jenis validitas yang biasa digunakan dalam mengukur validitas instrumen penelitian?
3. Jelaskan apa yang dimaksud dengan reliabilitas?
4. Apa perbedaan antara validitas dan reliabilitas?

DAFTAR PUSTAKA

- Sudjana. 2005. Metoda statistika. Tarsito. Bandung
- Riwidikdo, H. 2012. Statistik kesehatan. Mitra cendikia press. Yogyakarta
- Budiarto, E. 2001. Biostatistika untuk kedokteran dan kesehatan masyarakat. EGC. Bandung
- Mason, R.D & Douglas A. Lind. 1996. Teknik statistik bisnis dan ekonomi. Erlangga. Jakarta
- Usman, H. 2000. Pengantar statistika. Bumi aksara. Bandung

9

POKOK BAHASAN UJI PARAMETRIK DAN NON PARAMETRIK

9.1. PENDAHULUAN

9.1.1. Deskripsi Singkat

Uji yang akan dibahas dalam bab ini adalah uji parametrik dan nonparametrik. Kegunaan uji parametrik dan nonparametrik. Ketentuan dalam pemilihan uji parametrik dan non parametrik.

9.1.2. Relevansi

Materi yang disajikan pada pokok bahasan ini meliputi pengertian statistik parametrik dan nonparametrik, syarat uji dan macam uji parametrik dan nonparametrik.

9.1.3. Kompetensi

Pada akhir perkuliahan mahasiswa diharapkan mampu:

1. Menjelaskan statistik parametrik dan non parametrik
2. Menjelaskan syarat uji statistik parametrik dan non parametrik
3. Membedakan penggunaan uji statistik parametrik dan non parametrik

9.2. PENJELASAN

9.2.1. Uraian Isi

A. Pengertian statistik parametrik dan non parametrik

Secara garis besar ilmu statistik dibagi menjadi dua yaitu:

- a. Statistik parametrik
- b. Statistik non parametrik

Statistik parametrik yaitu statistik yang digunakan untuk menguji hipotesis yang variabelnya terukur, digunakan untuk menguji data dengan sebaran data normal dengan skala data yang digunakan adalah interval dan rasio. Contoh uji statistik parametrik yaitu uji-z, uji-t, korelasi pearson, anova dan lain sebagainya. Adapun keunggulan statistik parametrik adalah hasil uji statistik parametrik terhadap populasi tajam.

Statistik non parametrik yaitu statistik yang digunakan untuk menguji hipotesis yang variabelnya tidak memiliki kepastian (standart), digunakan untuk menguji data tanpa melihat sebaran data baik normal maupun tidak normal dengan skala data yang digunakan adalah nominal dan ordinal. Contoh metode statistik yang digunakan dalam statistik non parametrik yaitu Rank spearman, fisher exact, chi square dan lain sebagainya. Adapun keunggulan statistik non parametrik adalah tidak membutuhkan asumsi normalitas dan jumlah sampel kecil.

B. Syarat uji statistik parametrik dan non parametrik

Dalam mengolah data diperlukan kejelian dalam pemilihan uji statistik yang akan digunakan. Berikut

beberapa syarat yang harus diperhatikan dalam memilih uji:

- a. Syarat uji statistik parametrik
 1. Distribusi data normal
 2. Sampel diambil secara random
 3. Varians kelompok sama
 4. Skala pengukuran interval/rasio
- b. Syarat uji statistik nonparametrik
 1. Distribusi data normal maupun tidak normal
 2. Sampel diambil secara random maupun tidak
 3. Variasi kelompok tidak sama
 4. Skala pengukuran nominal/ordinal

Jika salah satu syarat uji statistik parametrik tidak dapat dipenuhi maka uji statistik harus menggunakan uji statistik nonparametrik.

C. Uji statistik parametrik

Keunggulan uji statistik parametrik:

- a. Syarat parameter dari suatu populasi yang menjadi sampel biasanya tidak diuji dan dianggap memenuhi syarat, pengukuran terhadap data dilakukan dengan kuat
- b. Observasi bebas satu sama lain dan ditarik dari populasi yang berdistribusi normal serta memiliki varian yang homogen.

Macam-macam uji statistik parametrik:

- a. z-test
- b. t-test
- c. tes proporsi
- d. uji korelasi pearson
- e. analisis varian

D. Uji statistik non parametrik

Keunggulan statistik non parametrik:

- a. Asumsi uji nonparametrik lebih longgar, yaitu jika syarat uji parametrik tidak dapat terpenuhi (misalnya distribusi data tidak normal) maka statistik nonparametrik lebih sesuai digunakan
- b. Perhitungannya dapat dilaksanakan dengan cepat dan mudah
- c. Uji statistik nonparametrik dapat diterapkan jika menghadapi keterbatasan data misalnya skala data lemah (nominal/ordinal)
- d. Dengan jumlah sampel yang sedikit uji statistik nonparametrik lebih efisien.

Macam-macam uji statistik nonparametrik:

- a. Uji tanda
- b. Uji wilcoxon
- c. Uji rank spearman
- d. Uji kendall
- e. Uji run
- f. Uji median
- g. Uji chi square

E. Menentukan Uji Statistik

Untuk memudahkan memilih uji statistik yang tepat untuk diaplikasikan kedalam penelitian, dapat dilihat tabel dibawah ini

Tabel. Jenis Uji Statistik

Macam Data	Bentuk Hipotesis				
	Komparatif (2 sampel)		Komparatif (>2 sampel)		Asosiasi (Hubungan)
	Related	Independen	Related	Independen	
Nominal	Mc Nemar	<ul style="list-style-type: none"> Fisher exact-probability Chi-Square 	<ul style="list-style-type: none"> X² untuk K sampel Cochran Q 	<ul style="list-style-type: none"> X² untuk K sampel 	<ul style="list-style-type: none"> Contingency Coefficient C
Ordinal	<ul style="list-style-type: none"> Sign Test Wilcoxon matched pairs 	<ul style="list-style-type: none"> Median test Mann-whitney U test Kolmogorof-smirnov Wald-woldfowitz 	<ul style="list-style-type: none"> Friedman-two-way-anova 	<ul style="list-style-type: none"> Median extension Kruskal wallis-one-way-anova 	<ul style="list-style-type: none"> Spearman rank correlation Kendall tau
Interval, rasio	<ul style="list-style-type: none"> T-test of related (pired) 	<ul style="list-style-type: none"> T-test independent 	<ul style="list-style-type: none"> One way anova Two way anova 	<ul style="list-style-type: none"> One way anova Two way anova 	<ul style="list-style-type: none"> Pearson Pruduct Momment Partial Correlation Multiple Correlation Regresi

Sumber: Sugiyono (1999). Statistik Nonparametrik Untuk Penelitian, CV Alfabeta, Bandung

Untuk mempermudah menentukan uji beda (komparatif) dapat dilihat atbel dibawah ini:

Tabel. Uji beda/komparatif

Jumlah variabel	Keterikatan variabel	Uji paramterik	Uji non-parametrik
2 variabel	Independen	Uji t-test (n kecil) Uji Z (n besar)	Mann-whitney/uji median Uji chi-square
	Related	Paired t-test	Sign-test Wilcoxon-test Uji mc-nemar
>2 variabel	Independen	Anova/Uji F Manova (Multivariat Anova)	Kruskal wallis
	Related	Repeated measure	Friedman Kendall's w Cohran's Q

9.3. PENUTUP

Latihan

1. Jelaskan perbedaan uji statistik parametrik dan nonparametrik!
2. Sebutkan syarat yang harus dipenuhi untuk dapat menggunakan uji statistik parametrik?
3. Jika syarat uji parametrik tidak terpenuhi apa yang harus dilakukan peneliti?

DAFTAR PUSTAKA

Sudjana. 2005. Metoda statistika. Tarsito. Bandung

Riwidikdo, H. 2012. Statistik kesehatan. Mitra cendikia press. Yogyakarta

Budiarto, E. 2001. Biostatistika untuk kedokteran dan kesehatan masyarakat. EGC. Bandung

Mason, R.D & Douglas A. Lind. 1996. Teknik statistik bisnis dan ekonomi. Erlangga. Jakarta

Usman, H. 2000. Pengantar statistika. Bumi aksara. Bandung

10

POKOK BAHASAN UJI BEDA (T-test dan Anova)

10.1. PENDAHULUAN

10.1.1. Deskripsi Singkat

Uji yang akan dibahas dalam bab ini adalah uji beda yang terdiri dari uji T dan uji Anova.

10.1.2. Relevansi

Materi yang disajikan pada pokok bahasan ini meliputi uji beda t-test one sample, beda T-test two sampel berpasangan, uji beda T-test two sampel tidak berpasangan, uji beda nonparametrik, uji beda one way anova, dan uji beda two way anova.

10.1.2. Kompetensi:

Pada akhir perkuliahan mahasiswa diharapkan mampu:

1. Menjelaskan uji beda T-test one sampel
2. Menjelaskan uji beda T-test two sampel berpasangan
3. Menjelaskan uji beda T-test two sampel tidak berpasangan

4. Menjelaskan uji beda nonparametrik
5. Menjelaskan uji beda one way anova
6. Menjelaskan uji beda two way anova

10.2. PENJELASAN

10.2.1. Uraian Isi

A. Uji beda T-test one sampel

Uji beda one sampel T-test merupakan salah satu jenis uji statistik parametrik yang bertujuan untuk mengetahui apakah rata-rata suatu sampel secara statistik berbeda atau sama dengan suatu angka. Agar uji ini dapat digunakan terdapat beberapa syarat yang harus dipenuhi, yaitu sampel diambil secara acak, data berdistribusi normal dan skala data interval/rasio. Rumus yang digunakan untuk uji t-test one sampel sebagai berikut:

$$t \text{ stat} = \frac{X - \mu_0}{\frac{s}{\sqrt{n}}}$$

Dimana:

X = rata-rata sampel

μ_0 = rata-rata standar

s = standar deviasi

n = jumlah sampel

Contoh aplikasi uji beda T-test one sampel

1. Kadar Hb standar normal tidak anemia dipergunakan angka 11. Berdasarkan penelitian di lapangan terhadap ibu-ibu pekerja pertanian didapatkan rata-rata kadar Hb 10,8 dengan standar deviasi 0,5 dari pengujian 30 sampel ibu. Selidikilah dengan $\alpha = 1\%$, apakah kadar Hb ibu-ibu pekerja pertanian di bawah standar normal tidak anemia ?

Penyelesaian:

Langkah 1:

- Hipotesis
 - ✓ $H_0 : Hb_{10,8} = Hb_{11}$; tidak ada beda kadar Hb ibu-ibu pekerja pertanian dengan standar normal tidak anemia
 - ✓ $H_a : Hb_{10,8} < Hb_{11}$; ada beda kadar Hb ibu-ibu pekerja pertanian dengan standar normal tidak anemia
- Level signifikansi (α)
 - ✓ $\alpha = 1\%$

Langkah 2:

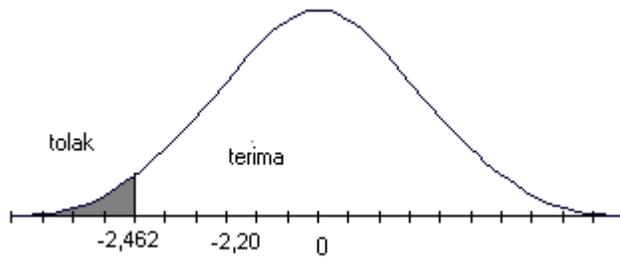
Hitung rumus statistik penguji:

- $X=10,8$; $\mu_0=11$; $SD=0,5$; $N=30$

$$t = \frac{\bar{X} - \mu_0}{\frac{SD}{\sqrt{N}}}$$
$$t = \frac{10,8 - 11}{\frac{0,5}{\sqrt{30}}}$$
$$t = -2,20$$

Langkah 3:

- Df/db/dk
 - ✓ $Df = N - 1 = 30 - 1 = 29$
- Nilai tabel
 - ✓ Nilai tabel t distribusi student. Uji satu sisi, $\alpha = 1\%$, $df = 29$, nilai t tabel = 2,462
- Daerah penolakan



- ✓ $|-2,20| < 2,462$;
- ✓ berarti H_0 diterima
- ✓ H_a ditolak
- ✓ Kesimpulan
- ✓ Tidak ada beda kadar Hb ibu-ibu pekerja pertanian dengan standar normal tidak anemia pada $\alpha=1\%$.

B. Uji beda T-test two sampel berpasangan

Dua sampel berpasangan disini diartikan sebagai sebuah sampel dengan subjek yang sama namun mengalami dua perlakuan atau pengukuran yang berbeda. Uji ini bertujuan untuk mengetahui apakah dua sampel berpasangan mempunyai nilai rata-rat yang sama atau tidak. Adapun syarat yang harus dipenuhi dalam melakukan uji ini adalah:

- a. Satu sampel (setiap elemen memiliki 2 nilai pengamatan, sebelum-sesudah)
- b. Skala data interval/rasio
- c. Distribusi data normal

Rumus yang digunakan sebagai berikut:

$$t = \frac{\bar{X}_D - \mu_0}{s_D / \sqrt{n}}$$

Dimana:

$$\bar{X}_d = \frac{\sum D}{n}$$
$$s_d = \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}}$$

Dengan: D = selisih X1 dan X2 (X1-X2)

n = jumlah sampel

X = rata-rata

SD = Standart deviasi dari D

Langkah-langkah pengujian :

1. Tetapkan Ho dan H1
2. Tetapkan titik kritis yang terdapat pada tabel 't'
3. Tentukan daerah kritis dengan db=n-1
4. Tentukan t hitung dengan menggunakan rumus
5. Lakukan uji signifikansi dengan membandingkan t hitung dengan t tabel

Contoh kasus:

1. Suatu kegiatan penelitian eksperimental, telah berhasil menemukan metode "ABG" sebagai metode baru untuk mengajarkan mata kuliah Statistik II. Dalam rangka uji coba terhadap efektifitas atau kemampuan metode baru itu, dilaksanakan penelitian lanjutan dengan mengajukan Hipotesis Nol (Nihil) yang mengatakan : Tidak terdapat perbedaan yang signifikan nilai Statistik II antara sebelum dan sesudah di terapkannya metode "ABG" sebagai metode mengajar. Dalam rangka pengujian ini diambil sampel sebanyak 20 mahasiswa. Gunakan taraf kepercayaan 95 % (alfa=5%) untuk menguji pernyataan (Hipotesis) tersebut. Dengan data sebagai berikut:

Nama	Nilai Statistik II	
	Sebelum	Sesudah
A	78	75
B	60	68
C	65	59
D	55	71
E	70	63
F	49	54
G	68	66
H	70	74
I	81	89
J	30	33
K	55	51
L	40	50
M	63	68
N	85	83
O	70	77
P	62	69
Q	58	73
R	65	65
S	75	76
T	69	86

Penyelesaian:

1. Menentukan hipotesis:

Ho = tidak ada perbedaan yang signifikan antara hasil belajar sebelum dan sesudah menggunakan metode ABG

Ha = ada perbedaan yang signifikan antara hasil belajar sebelum dan sesudah menggunakan metode ABG

2. Tetapkan titik kritis yaitu alfa 5%
3. Tentukan daerah kritis $D_b = n - 1 = 20 - 1 = 19$
4. Tentukan t hitung

Sebelum (x_1)	Sesudah (x_2)	D= x_1-x_2	D ²
78	75	3	9
60	68	-8	64
55	59	-4	16
70	71	-1	1
57	63	-6	36
49	54	-5	25
68	66	2	4
70	74	-4	16
81	89	-8	64
30	33	-3	9
55	51	4	16
40	50	-10	100
63	68	-5	25
85	83	2	4
70	77	-7	49
62	69	-7	49
58	73	-15	225
65	65	0	0
75	76	-1	1
69	86	-17	289
Jumlah		-90	1002

Menghitung Sd =

$$\begin{aligned}
 s &= \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}} \\
 &= \sqrt{\frac{1}{19} \left\{ 1002 - \frac{(-90)^2}{20} \right\}} \\
 &= \sqrt{31,4211} = 5,6054
 \end{aligned}$$

Menghitung t

$$t = \frac{-90}{\frac{20}{5.6054}} = \frac{-4,50}{1.2534} = -3,5902 .$$

5. Lakukan uji signifikansi
T tabel = 2,093 maka, t hitung > t tabel
Sehingga disimpulkan Ho ditolak.

C. Uji beda T-two sampel tidak berpasangan

Uji ini digunakan untuk menguji data yang saling independent (tidak berpasangan) tujuannya yaitu untuk mengetahui ada atau tidaknya perbedaan rerata antara dua buah data.

Adapun syarat yang harus dipenuhi yaitu:

- Data berdistribusi normal
- Data dipilih secara acak
- Data yang digunakan dengan skala interval/rasio

Dalam pengujian t-test two sampel saling bebas ada 4 cara yang digunakan untuk memperoleh t-hitung, yaitu

- Jika varian populasi diketahui

Cara ini digunakan jika nilai varian populasi diketahui, yaitu dengan mencari nilai z sebagai berikut:

$$z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_{X_1}^2}{n_1} + \frac{\sigma_{X_2}^2}{n_2}}} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- Varian populasi tidak diketahui, ukuran sampel sama dan varian diasumsikan sama

Digunakan jika ukuran sampel (n) sama dan varians dianggap sama. Dengan menghitung nilai t sebagai berikut:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1 X_2} \cdot \sqrt{\frac{2}{n}}}$$

Dimana

$$s_{X_1 X_2} = \sqrt{\frac{1}{2}(s_{X_1}^2 + s_{X_2}^2)}$$

- Varian populasi tidak diketahui, ukuran sampel berbeda dan varians diasumsikan sama

Meskipun varian homogen namun ukuran sampel yang digunakan berbeda maka menggunakan rumus t-hitung berikut;

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1X_2} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Dimana

$$s_{X_1X_2} = \sqrt{\frac{(n_1 - 1)s_{X_1}^2 + (n_2 - 1)s_{X_2}^2}{n_1 + n_2 - 2}}$$

- d. Varian populasi tidak diketahui, ukuran sampel sama/ berbeda, varian diasumsikan berbeda

Uji ini disebut juga sebagai welch's test dan hanya digunakan jika varian diasumsikan berbeda (baik ukuran sampel sama atau berbeda). Yaitu menggunakan rumus sebagai berikut;

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}$$

Dimana

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Untuk menentukan *degree of freedom* menggunakan rumus sebagai berikut;

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

Persamaan ini disebut sebagai **persamaan welch satterthwaite**

Langkah-langkh uji:

1. Pastikan data dipilih secara acak dan berdistribusi normal
2. Tentukan apakah varian homogen atau heterogen
3. Tuliskan H_a dan H_o
4. Cari t-hitung atau z-hitung berdasarkan rumus yang sesuai
5. Tentukan taraf signifikan
6. Cari t tabel atau z tabel dimana df tergantung pada rumus yang dipilih
7. Tentukan kriteria pengujian kapan H_o diterima atau ditolak
8. Bandingkan t-hitung dengan t-tabel
9. Buatlah kesimpulan

D. Uji beda nonparametrik

Jika uji beda diatas tidak dapat diaplikasikan maka dapat menggunakan alternatif uji beda nonparamterik yaitu uji mann-whitney dan uji wilcoxon.

1. Uji mann-whitney (uji beda dua kelompok independen)

Uji ini dapat digunakan jika jumlah sampel sedikit yaitu kurang dari 30 sampel. Data tidak berdistribusi normal dan digunakan untuk menguji data dengan skala data ordinal, interval, rasio. Uji ini dapat digunakan sebagai alternatif ketika t-test tidak dapat digunakan. Syarat uji ini yaitu:

- a. Data berskala ordinal, interval atau rasio
- b. Terdiri dari dua kelompok independent atau saling bebas
- c. Kelompok 1 dan 2 jumlah datanya tidak harus sama banyak

d. Data tidak harus berdistribusi normal.

Prosedur pengujian:

1. Susun kedua hasil pengamatan menjadi satu kelompok sampel
2. Hitung jenjang/rangking untuk tiap nilai dalam sampel gabungan
3. Jenjang atau rangking diberikan mulai nilai terkecil hingga terbesar
4. Nilai beda sama diberi jenjang nilai rata-rata
5. Jumlahkan nilai jenjang masing-masing sampel
6. Hitung nilai U

Ada 2 macam untuk menghitung nilai U yaitu:

a. Untuk sampel kecil (≤ 20)

$$U_1 = n_1 \cdot n_2 - U_2$$

$$U_2 = n_1 \cdot n_2 - U_1$$

Dengan nilai U 1 dan U2 diperoleh dari:

$$U_1 = n_1 \cdot n_2 + \frac{n_2(n_2 + 1)}{2} - \sum R_2$$

Dengan: U1 = statistik uji U1

U2 = statistik uji U2

R1 = jumlah rank sampel 1

R2 = jumlah rank sampel 2

n1 = banyaknya sampel 1

n2 = banyaknya sampel 2

Setelah mendapatkan nilai statistik uji U1 dan U2. kemudian mengambil nilai terkecil dari kedua nilai tersebut. Nilai terkecil yang diperoleh kemudian dibandingkan dengan tabel mann whitney.

b. Sampel besar (n_1 atau $n_2 > 20$)

- Untuk ranking yang sama

$$Z = \frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}}$$

- Untuk ranking yang berbeda

$$Z = \frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\left(\frac{n_1 \cdot n_2}{(n_1 + n_2) \cdot (n_1 + n_2 - 1)}\right) \left(\frac{(n_1 + n_2)^3 - (n_1 + n_2)}{12} - \sum \frac{t_i^3 - t_i}{12}\right)}}$$

Contoh perhitungan:

1. Seorang peneliti ingin mengetahui perbedaan denyut nadi antara laki-laki dan perempuan. Diperoleh data sebagai berikut:

Laki-laki	Perempuan
90	79
89	82
82	85
89	88
91	85
86	80
85	80
86	
84	

Penyelesaian:

1. Tentukan hipotesis:

Ho = denyut nadi wanita sama dengan denyut nadi pria

Ha = denyut nadi wanita berbeda dengan denyut nadi pria

2. Susun kedua hasil pengamatan menjadi satu kelompok sampel

Denyut nadi	Ranking	Jenis kelamin
79	1	Wanita
80	2,5	Wanita
80	2,5	Wanita
82	4,5	Pria
82	4,5	Wanita
84	6	Pria
85	8	Pria
85	8	Wanita
85	8	Wanita
86	10,5	Pria
86	10,5	Pria
88	12	Wanita
89	13,5	Pria
89	13,5	Pria
90	15	Pria
91	16	Pria

3. Jumlahkan ranking

Denyut Nadi Pria	Rangking	Denyut Nadi Wanita	Rangking
90	15	79	1
89	13,5	82	4,5
82	4,5	85	8
89	13,5	88	12
91	16	85	8
86	10,5	80	2,5
85	8	80	2,5
86	10,5		
84	6		
Jumlah Rangking	97,5		38,5

4. Hitung nilai U

$$U_1 = n_1 \cdot n_2 + \frac{n_2(n_2 + 1)}{2} - \sum R_2$$

$$U_1 = 9 \cdot 7 + \frac{7(7+1)}{2} - 38,5 = 52,5$$

$$U_2 = n_1 \cdot n_2 - U_1 \\ = 10,5$$

5. Nilai U terkecil adalah 10,5 maka bandingkan nilai U terkecil dengan U tabel
6. Kesimpulan H_0 ditolak

2. Uji Wilcoxon

Syarat uji wilcoxon adalah sebagai berikut:

- Menggunakan data berpasangan dari populasi yang sama
- Sampel dipilih secara acak
- Skala data minimal ordinal
- Tidak memperhatikan normalitas data

Langkah-langkah pengujian wilcoxon:

- Berikan jenjang (rank) untuk tiap beda dari pasangan pengamatan ($y_i - x_i$) sesuai dengan besarnya, dari yang terkecil sampai terbesar tanpa memperhatikan tanda dari beda itu (nilai beda absolut).
- Bila ada dua atau lebih beda yang sama, maka jenjang untuk tiap-tiap beda itu adalah jenjang rata-rata
- Bubuhkan tanda positif atau negatif pada jenjang untuk tiap beda sesuai dengan tanda dari beda itu. Beda 0 tidak diperhatikan
- Jumlahkan semua jenjang bertanda positif atau negatif, tergantung dari mana yang memberikan jumlah yang lebih kecil setelah tandanya

dihilangkan. Notasi jumlah jenjang yang lebih kecil ini dengan T

5. Bandingkan nilai T yang diperoleh dengan nilai t uji wilcoxon

Contoh penggunaan uji wilcoxon

Seorang dokter ingin melakukan penelitian ingin melihat pengaruh dari suatu obat. Delapan orang pasien yang diambil secara acak diukur kapasitas pernapasannya sebelum dan sesudah diberikan obat tertentu. Hasilnya sebagai berikut :

Pasien	A	B	C	D	E	F	G	H
Sebelum	2750	2360	2950	2830	2250	2680	2720	2810
Sesudah	2850	2380	2930	2860	2300	2640	2760	2800

dengan menggunakan $\alpha = 0,05$

Penyelesaian:

1. Hipotesis:

H_0 : Tidak ada perbedaan sebelum dan sesudah menggunakan obat

H_1 : Ada perbedaan sebelum dan sesudah menggunakan obat

2. Taraf nyata dan nilai T tabelnya

$\alpha = 0,05$ dengan $n = 8$ tabel wilcoxon $T = 3$.
(diperoleh dari tabel wilcoxon)

3. Kriteria Pengujian

H_0 diterima apabila nilai uji statistik \geq dari t tabel yaitu 3.

H_0 ditolak apabila nilai uji statistik $<$ dari t tabel yaitu 3.

4. Nilai uji statistik

Pasien	Sebelum	Sesudah	selisih (d)	Peringkat
A	2750	2850	-100	-8
B	2360	2380	-20	-2,5
C	2950	2930	20	2,5
D	2830	2860	-30	-4
E	2250	2300	-50	-7
F	2680	2640	40	5,5
G	2720	2760	-40	-5,5
H	2810	2800	10	1

Menjumlahkan nilai berdasarkan tanda.

Untuk tanda positif: $2,5 + 5,5 + 1 = 9$

Untuk tanda negatif $8 + 2,5 + 4 + 7 + 5,5 = 27$

Untuk melihat nilai uji statistiknya yaitu dari nilai terkecil dari nilai tersebut yaitu tanda positif 9. sehingga nilai statistiknya 9.

5. Kesimpulan

Berdasarkan hasil tersebut diperoleh hasil bahwa nilai uji statistik \geq dari t tabel. yaitu $9 \geq 3$. sehingga berdasarkan kriteria pengujian diperoleh hasil terima H_0 . sehingga disimpulkan bahwa tidak ada perbedaan sebelum dan sesudah menggunakan obat.

E. Anova

Analisis of varians (ANOVA) uji anova digunakan jika ingin menguji rata-rata kelompok lebih dari atau sama dengan tiga. Dalam analisis anova hanya digunakan hipotesis dua arah yaitu apakah ada perbedaan rata-rata.

Berikut penjelasan mengapa menggunakan anova:

1. Mengetahui signifikansi perbedaan rata-rata (μ) antara kelompok sampel yang satu dengan yang lain.
2. Analisis varians relatif mudah dimodifikasi dan dapat dikembangkan untuk berbagai bentuk percobaan yang lebih rumit. Selain itu, analisis ini juga masih memiliki keterkaitan dengan analisis regresi. Akibatnya, penggunaannya sangat luas di berbagai bidang, mulai dari eksperimen laboratorium hingga eksperimen periklanan, psikologi, dan kemasyarakatan.

Adapun syarat yang harus dipenuhi oleh uji anova yaitu:

1. Data berdistribusi normal, karena pengujiannya menggunakan uji F-Snedecor
2. Varians atau ragamnya homogen, dikenal sebagai homoskedastisitas, karena hanya digunakan satu penduga (estimate) untuk varians dalam contoh
3. Masing-masing contoh saling bebas, yang harus dapat diatur dengan perancangan percobaan yang tepat
4. Komponen-komponen dalam modelnya bersifat aditif (saling menjumlah).

Anova sendiri terbagi menjadi dua yaitu one way anova dan two way anova.

5. One way anova

Untuk menguji perbedaan rata-rata lebih dari dua sampel dimana dalam melakukan analisis hanya bisa satu arah. Maksud satu arah ini hanya bisa menguji antar kelompok yang satu. Untuk lebih jelasnya lihat contoh kasus berikut :

Contoh kasus Anova satu arah:

Sampel	Penurunan Berat Badan (Kg)			
	Metode 1	Metode 2	Metode 3	Metode 4
Sampel 1	4	8	7	6
Sampel 2	6	12	3	5
Sampel 3	4	-	-	5

Terdapat 4 metode diet dan 3 golongan usia peserta program diet. Berikut data rata-rata penurunan berat peserta keempat metode dalam tiga kelompok umur. Berdasarkan gambar di atas terlihat bahwa ada empat metode (kolom). Dari empat metode itu dilakukan oleh beberapa orang tapi tiap metode dilakukan oleh orang yang berbeda. Pada tabel di atas terlihat data diperoleh dari sampel yang berbeda perlakuan antar kelompok karena itu kita hanya bisa membandingkan antar metode tapi tidak bisa membandingkan antar orang karena setiap orang tidak melakukan metode yang sama oleh karena itu dikatakan satu arah saja.

6. Two way anova
7. Anova dua arah tanpa interaksi anova two way without interaction

Jenis anova yang kedua yaitu anova dua arah tanpa interaksi. Artinya bahwa bisa dilakukan **interaksi antara kelompok dan perlakuan**. Maksudnya bisa membandingkan antar antar kelompok atau kah antar perlakuan. berikut contoh kasus.

Contoh kasus Anova dua arah tanpa interaksi:

Umur	Penurunan Berat Badan (Kg)			
	Metode 1	Metode 2	Metode 3	Metode 4
< 20 tahun	5	6	2	3
20-40	2	7	5	3
> 40 tahun	7	3	4	3

Terdapat 4 metode diet dan 3 golongan usia peserta program diet Berikut data rata-rata penurunan berat peserta keempat metode dalam tiga kelompok umur. Berdasarkan gambat tersebut terlihat bahwa setiap metode memiliki perlakuan yang sama sehingga bisa dikatakan ada hubungan dua arah. Tapi tidak ada interaksi.

8. Anova dua arah dengan interaksi anova two way with interaction

Dikatakan anova dengan interaksi ketika setiap kolom (perlakuan) dan blok (baris) diulang.

Contoh kasus Anova dua arah dengan interaksi:

Umur	Penurunan Berat Badan (Kg)			
	Metode 1	Metode 2	Metode 3	Metode 4
< 20 tahun				
#1	5	0	3	4
#2	4	2	4	2
#3	5	1	8	2
20-40 tahun				
#1	5	4	2	5
#2	6	2	2	3
#3	2	1	4	2
> 40 tahun				
#1	4	5	2	6
#2	4	5	1	4
#3	5	0	2	4

Terdapat 4 metode diet, 3 kelompok umur dan 3 ulangan. Berikut adalah data rata-rata penurunan berat badan setelah 1 bulan melakukan diet. Ujilah apakah penurunan berat badan sama untuk setiap metode diet, kelompok umur dan interaksi dengan taraf uji 5 %?

Rumus uji Anova adalah sebagai berikut :

$$F = \frac{Sb^2}{Sw^2}$$

DF = Numerator (pembilang) = k-1, Denominator (penyebut) = n-k

Dimana varian between :

$$Sb^2 = \frac{n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + \dots + n_n(\bar{x}_n - \bar{x})^2}{k - 1}$$

Dimana rata-rata gabungannya :

$$\bar{x} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2 + \dots + n_n \cdot \bar{x}_n}{k - 1}$$

Sementara varian within :

$$Sw^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + \dots + (n_n - 1)S_n^2}{n - k}$$

KETERANGAN :

Sb = varian between

Sw = varian within

Sn² = varian kelompok

X = rata-rata gabungan

Xn = rata-rata kelompok

Nn = banyaknya sampel pada kelompok

k = banyaknya kelompok

10.3. PENUTUP

Latihan

1. sebuah penelitian tentang perbedaan kemampuan siswa dalam mempelajari matematika yang didasarkan pada pengelompokan siswa menurut jenis pekerjaan orang tua, diperoleh data sebagai berikut:

Peg. Negeri dan ABRI	Guru	Tani	Buruh tani
70	80	75	65
60	85	80	70
75	75	85	75
65	90	80	60
65	85	70	60
90	75	75	75
85	85	70	75
90	95	60	75
75	90	75	70
70	100	95	65

Apakah hasil belajar siswa dalam bidang matematika berbeda secara signifikan antara kelompok diatas?

DAFTAR PUSTAKA

- Sudjana. 2005. Metoda statistika. Tarsito. Bandung
- Riwidikdo, H. 2012. Statistik kesehatan. Mitra cendikia press. Yogyakarta
- Budiarto, E. 2001. Biostatistika untuk kedokteran dan kesehatan masyarakat. EGC. Bandung
- Mason, R.D & Douglas A. Lind. 1996. Teknik statistik bisnis dan ekonomi. Erlangga. Jakarta
- Usman, H. 2000. Pengantar statistika. Bumi aksara. Bandung

11

POKOK BAHASAN UJI KORELASI

11.1. PENDAHULUAN

11.1.1. Deskripsi Singkat

Uji yang akan dibahas dalam bab ini adalah uji beda korelasi Pearson Product Moment dan Rank Spearman.

11.1.2. Relevansi

Materi yang disajikan pada pokok bahasan ini meliputi uji korelasi Product Moment dan Rank Spearman. Langkah-langkah uji Product Moment dan Rank Spearman.

11.1.3. Kompetensi:

Pada akhir perkuliahan ini mahasiswa diharapkan mampu:

1. Menjelaskan uji korelasi Pearson Product Moment
2. Menjelaskan uji korelasi Rank Spearman

11.2. PENJELASAN

11.2.1. Uraian Isi

A. Uji Korelasi Product Moment

Kegunaan uji ini adalah untuk menyatakan ada atau tidaknya hubungan antara variabel X dengan variabel Y. Untuk menyatakan besarnya sumbangan variabel satu terhadap yang lainnya yang dinyatakan dalam persen. Dengan syarat sebagai berikut:

- Data berdistribusi Normal
- Variabel yang dihubungkan mempunyai data linear.
- Variabel yang dihubungkan mempunyai data yang dipilih secara acak.
- Variabel yang dihubungkan mempunyai pasangan yang sama dari subyek yang sama pula (variasi skor variabel yang dihubungkan harus sama).
- Variabel yang dihubungkan mempunyai data interval atau rasio.
- Nilai r terbesar adalah $+1$ dan r terkecil adalah -1 . $r = +1$ menunjukkan hubungan positif sempurna, sedangkan $r = -1$ menunjukkan hubungan negatif sempurna.
- r tidak mempunyai satuan atau dimensi. Tanda $+$ atau $-$ hanya menunjukkan arah hubungan. Interpretasi nilai r adalah sebagai berikut:

r	Interpretasi
0	Tidak berkorelasi
0,01-0,20	Korelasi Sangat rendah
0,21-0,40	Rendah
0,41-0,60	Agak rendah
0,61-0,80	Cukup
0,81-0,99	Tinggi
1	Sangat tinggi

B. Langkah-langkah Menghitung Koefisien Korelasi Parsial

- a. Tulis Ho dan Ha dalam bentuk kalimat.
- b. Tulis Ho dan Ha dalam bentuk statistik.
- c. Buat tabel penolong sebagai berikut:

No. resp.	X	Y	XY	x ²	Y ²

- d. Cari r_{hitung}.

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

- e. Tentukan taraf signifikansinya (α)
- f. Cari r tabel dengan $dk = n-2$
- g. Tentukan kriteria pengujian
Jika $-r_{tabel} \leq r_{hitung} \leq +r_{tabel}$, maka Ho diterima
- h. Bandingkan t_{hitung} dengan t_{tabel}
- i. Buatlah kesimpulan.

Contoh:

- 1. Tulis Ho dan Ha dalam bentuk kalimat.
Ho : Tidak terdapat hubungan yang positif dan signifikan antara variabel Biaya Promosi dengan Nilai Penjualan.
Ha : Terdapat hubungan yang positif dan signifikan antara variabel Biaya Promosi dengan Nilai Penjualan.
- 2. Tulis Ho dan Ha dalam bentuk statistik. Ho : $r = 0$.
Ha : $r \neq 0$.
- 3. Buat tabel penolong sebagai berikut:

Nilai Penjualan Y	Biaya Promosi X	XY	X ²	Y ²
64	20	1280	400	4096
61	16	976	256	3721
84	34	2856	1156	7056
70	23	1610	529	4900
88	27	2376	729	7744
92	32	2944	1024	8464
72	18	1296	324	5184
77	22	1694	484	5929
$\Sigma Y = 608$	$\Sigma X = 192$	$\Sigma XY = 15032$	$\Sigma X^2 = 4902$	$\Sigma Y^2 = 47094$

4. Cari r_{hitung} .

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

$$= \frac{8(15.032) - (192)(608)}{\sqrt{8(4.902) - (192)^2} \sqrt{8(47.094) - (608)^2}}$$

Taraf signifikansi (α) = 0,05.

5. r_{tabel} dengan $dk = 8-2=6$ adalah 0,707

6. Tentukan kriteria pengujian

Jika $-r_{tabel} \leq r_{hitung} \leq +r_{tabel}$, maka H_0 diterima

7. Bandingkan r_{hitung} dengan r_{tabel}

r_{hitung} (0,86) > r_{tabel} (0,707), jadi H_0 ditolak.

8. Kesimpulan.

Terdapat hubungan yang positif dan signifikan antara variabel Biaya Promosi dengan Nilai Penjualan.

C. Uji korelasi rank spearman

Asumsi uji korelasi Spearman adalah:

- (1) Data tidak berdistribusi normal
- (2) Data diukur dalam skala Ordinal.

Rumus uji korelasi spearman untuk jumlah sampel ≤ 30 adalah:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Di mana:

r_s = Koefisien Korelasi Spearman

$\sum d^2$ = Total Kuadrat selisih antar ranking

n = Jumlah Sampel Penelitian

Langkah Uji:

1. Jumlahkan skor item-item di tiap variabel untuk mendapatkan skor total variabel (misalnya cari skor total variabel X dengan menotalkan item-item variabel X).
2. Lakukan rangking skor total x (r_x) dan rangking skor total y (r_y).
3. Cari nilai d yaitu selisih $r_x - r_y$.
4. Cari nilai d^2 yaitu kuadrat d (selisih $r_x - r_y$).

Contoh:

Responden	Item X #1	...	Item X #n	Total X	Item Y #1	...	Item Y #n	Total Y	r_x	r_y	d	D^2
1	5		4	65	5		4	57	1	1	0	0
2	4		4	62	4		4	55	4	3	1	1
3	4		3	61	3		3	50	5	6	-1	1
4	5		4	63	3		5	52	3	5	-2	4
5	4		5	64	2		4	54	2	4	-2	4
...												
30	4		5	60	5		3	56	6	2	4	16
total												2125

Setelah data dihitung dalam tabel, lalu masukkan ke dalam rumus uji korelasi Spearman:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6.2125}{30(30^2 - 1)} = 1 - \frac{12750}{30(900 - 1)} = 1 - \frac{12750}{30.799} = 0,47$$

Dengan demikian korelasi Spearman (r_s) variabel x dengan variabel y dalam contoh adalah 0,47. Nilai korelasi Spearman hitung ini (r_s) lalu diperbandingkan dengan Spearman Tabel (r_s tabel). Keputusan diambil dari perbandingan tersebut. Jika $r_s > r_s$ tabel, H_0 ditolak dan H_1 diterima. Jika r_s hitung $\leq r_s$ tabel, H_0 diterima, H_1 ditolak. Pengambilan keputusan dari contoh di atas adalah karena r_s hitung $> r_s$ tabel maka H_0 ditolak dan H_1 diterima. Artinya terdapat hubungan antara variabel x dengan variabel y .

Tabel Interpretasi Koefisien Korelasi Versi de Vaus

D.A. de Vaus menginterpretasikan koefisien korelasi sebagai berikut:

Koefisien	Kekuatan Hubungan
0,00	Tidak ada hubungan
0,01 – 0,09	Hubungan kurang berarti
0,10 – 0,29	Hubungan lemah
0,30 – 0,49	Hubungan moderat
0,50 – 0,69	Hubungan kuat
0,70 – 0,89	Hubungan sangat kuat
>0,90	Hubungan mendekati sempurna

Dalam contoh di atas maka kekuatan hubungan antara x dan y adalah hubungan moderat (karena 0,47).

Rumus di atas berlaku jika jumlah sampel lebih kecil atau sama dengan 30 (≤ 30). Cara menghitung uji korelasi Spearman dengan lebih dari 30 sampel dengan mencari Nilai z hitung terlebih dahulu.

Cara mencari nilai z hitung sebagai berikut:

$$z = r_s \sqrt{n - 1}$$

Di mana:

z = Nilai z hitung

r_s = Koefisien Korelasi Spearman

n = Jumlah Sampel Penelitian

Nilai rs dicari dengan cara yang sama seperti perhitungan terdahulu (di bagian atas). Dalam contoh sampel yang lebih besar dari 30 ini misalnya sampel menggunakan 50 responden. Maka perhitungannya sebagai berikut:

$$z = 0,99\sqrt{50 - 1} = 0,99\sqrt{49} = 0,99 \cdot 7 = 6,93$$

Nilai z hitung dalam sampel > 30 ini adalah 6,93. Pengambilan keputusan dalam sampel > 30 ini adalah membandingkan antara z hitung dengan z tabel. Z hitung sudah diperoleh sekarang tinggal z tabel.

Cara Mencari z Tabel

Nilai z tabel dicari dari tabel Z (lihat buku-buku statistik). Caranya adalah:

1. Tentukan Taraf Keyakinan Penelitian (misalnya 95%). Taraf Keyakinan 95% berarti Interval Keyakinan-nya (alpha) 0,05. Nilai 0,05 ini merupakan bentuk desimal dari 5% yang diperoleh dari pengurangan 100% selaku kebenaran absolut dengan 95% (100% - 95% = 5% atau 0,05).
2. Tentukan Uji yang digunakan. Apakah 1 sisi (One-Tailed) atau 2 sisi (Two-Tailed). Penentuan 1 sisi atau 2 sisi ini didasarkan hipotesis penelitian. Jika hipotesis hanya menyebutkan “terdapat hubungan” maka artinya bentuk hubungan belum ditentukan apakah positif atau negatif dan dengan demikian menggunakan uji 2 sisi. Jika hipotesis menyatakan “terdapat hubungan positif” atau “terdapat hubungan negatif” maka artinya bentuk hubungan sudah ditentukan dan dengan demikian menggunakan uji 1 sisi.

3. Jika Uji 2 Sisi (Two-Tailed) maka lihat Tabel Z. Dalam uji 2 sisi Interval Keyakinan dibagi dua yaitu $0,05 / 2 = 0,025$. Cari pada kolom tabel nilai yang paling mendekati 0,025. Dari nilai yang paling dekat tersebut tarik garis ke kiri sehingga bertemu dengan nilai $1,9 + 0,060 = 1,96$. Batas kiri pengambilan keputusan dengan kurva adalah $-1,96$ batas kanannya $+1,96$. Keputusannya: Tolak H_0 dan Terima H_1 jika $-z$ hitung $<$ dari $-1,96$ dan $>$ dari $+1,96$. Sebaliknya, Terima H_0 dan Tolak H_1 jika $-z$ hitung $>$ $-1,96$ dan $<$ dari $+1,96$.

11.3. PENUTUP

Latihan

1. Kapan rumus-rumus korelasi dibawah ini digunakan?
2. Korelasi pearson (*product momen correlation*)
3. Korelasi pearson dengan metode Z
4. Korelasi spearman
5. Sebuah penelitian yang mencari hubungan antara banyaknya jam belajar mandiri perminggu mahasiswa dengan hasil belajar (indeks prestasi mahasiswa). Dari 10 sampel yang diambil diperoleh data berikut:.

Mahasiswa	Jam belajar/minggu	IP yang dicapai
1	40	3,80
2	35	3,60
3	30	3,25
4	25	3,00
5	25	2,95
6	25	3,05
7	20	2,50
8	15	2,00
9	10	1,50
10	5	1,00

DAFTAR PUSTAKA

- Sudjana. 2005. Metoda statistika. Tarsito. Bandung
- Riwidikdo, H. 2012. Statistik kesehatan. Mitra cendikia press. Yogyakarta
- Budiarto, E. 2001. Biostatistika untuk kedokteran dan kesehatan masyarakat. EGC. Bandung
- Mason, R.D & Douglas A. Lind. 1996. Teknik statistik bisnis dan ekonomi. Erlangga. Jakarta
- Usman, H. 2000. Pengantar statistika. Bumi aksara. Bandung

12

POKOK BAHASAN CHI SQUARE DAN FISHER EXACT

12.1. PENDAHULUAN

12.1.1. Deskripsi Singkat

Uji yang akan dibahas dalam bab ini adalah uji chi square dan fisher exact.

12.1.2. Relevansi

Materi yang disajikan pada pokok bahasan ini meliputi uji chi square dan fisher exact.

12.1.3. Kompetensi

Diakhir pembelajaran mahasiswa diharapkan

1. Mampu menjelaskan chi square 2 sampel
2. Mampu menjelaskan fisher exact

12.2. PENJELASAN

12.2.1. Uraian Isi

A. Chi Square

Uji chi square sering disebut juga sebagai uji kai kuadrat. Uji ini merupakan salah satu uji statistik nonparametrik. Uji chi square berguna untuk menguji hubungan atau pengaruh dua buah variabel nominal dan mengukur kuatnya hubungan antara variabel satu dengan variabel nominal lainnya. (C=coefisien contingency)

Chi square mempunyai ketentuan sebagai berikut:

1. Nilai chi square tidak pernah negatif, karena selisih dari frekuensi pengamatan dan frekuensi harapan dikuadratkan.
2. Ketajaman dari distribusi chi square tidak tergantung pada ukuran sampel tetapi tergantung pada banyaknya kategori yang digunakan.
3. Distribusi chi square bersifat menceng kanan (nilai positif), semakin meningkat jumlah derajat bebas maka semakin mendekati distribusi normal.

Dengan rumus chi square

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

Dimana: χ^2 = nilai chi square

O = frekuensi yang diperoleh/diamati

E = Frekuensi yang diharapkan, dengan :

$$E = \frac{\text{Jumlah Sebaris} \times \text{Jumlah sekolom}}{\text{Jumlah total}}$$

Langkah-langkah/ prosedur dalam uji chi squared secara umum

1. Letakkan frekuensi-frekuensi terobservasi dalam k kategori. Jumlah frekuensi itu seluruhnya harus N, yakni banyak observasi-observasi independen.
2. Dari H₀ tentukan frekuensi yang diharapkan untuk tiap-tiap k sel itu. Manakala $k > 2$, dan bila lebih dari 20% dari E_i kurang dari 5, gabungkanlah kategori-kategori yang berdekatan apabila hal ini memungkinkan, dan dengan demikian kita mengurangi harga k serta meningkatkan nilai beberapa E_i. Apabila $k = 2$, tes X^2 untuk kasus satu sampel dapat digunakan secara memadai hanya jika tiap-tiap frekuensi yang diharapkan adalah lima atau lebih.
3. Hitung nilai X^2 dengan rumus $\sum(O_i - E_i)^2 / E_i$.
4. Tetapkan harga $df = k - 1$.
5. Dengan melihat tabel Chi squared, tetapkan probabilitas yang dikaitkan dengan terjadinya suatu harga yang sebesar nilai X^2 hitungan untuk harga df yang bersangkutan. Jika nilai ini sama atau kurang dari α , H₀ ditolak.

B. Fisher Exact

Merupakan salah satu uji nonparametrik yang digunakan untuk menganalisis dua sampel independen yang berskala nominal atau ordinal jika kedua sampel independennya berjumlah kecil (biasanya kurang dari 20). Data diklasifikasikan kedalam dua kelompok yang saling bebas sehingga akan terbentuk tabel kontingensi 2 x 2

12.3. PENUTUP

Latihan

1. Seorang pemilik pabrik berpendapat bahwa proporsi barang yang rusak, yang berasal dari 3 buah mesin yaitu mesin A, B, dan C adalah sama. Untuk menguji pendapat tersebut diambil 200 sampel acak yang terdiri dari 40 sampel produk mesin A, 40 sampel produk mesin B, dan 120 sampel produk mesin C. Ternyata 5 sampel produk mesin A rusak, 15 sampel produk mesin B rusak, dan 20 sampel produk mesin C rusak. Dengan menggunakan taraf nyata 5%, ujilah pendapat pemilik pabrik itu.
2. Suatu penelitian ingin mengetahui apakah ada perbedaan antara proporsi orang tua siswa di empat kota yaitu Jakarta, Bandung, Bogor dan Sukabumi yang setuju dengan penyuluhan tentang program X di sekolah menengah umum. Respons 500 orang tua siswa yang diambil secara acak dari masing-masing kota adalah sebagai berikut:

Kebiasaan Merokok	Jenis Kelamin	
	Setuju	Tidak Setuju
Jakarta	175	140
Bandung	80	50
Bogor	45	10
Sukabumi	40	70

Dengan menggunakan taraf nyata 5% ujilah data diatas.

DAFTAR PUSTAKA

- Sudjana. 2005. Metoda statistika. Tarsito. Bandung
- Riwidikdo, H. 2012. Statistik kesehatan. Mitra cendikia press. Yogyakarta
- Budiarto, E. 2001. Biostatistika untuk kedokteran dan kesehatan masyarakat. EGC. Bandung
- Mason, R.D & Douglas A. Lind. 1996. Teknik statistik bisnis dan ekonomi. Erlangga. Jakarta
- Usman, H. 2000. Pengantar statistika. Bumi aksara. Bandung

13

POKOK BAHASAN REGRESI

13.1. PENDAHULUAN

13.1.1. Deskripsi Singkat

Uji yang akan dibahas dalam pokok bahasan ini adalah uji regresi.

13.1.2. Relevansi

Materi yang disajikan pada pokok bahasan ini meliputi penggunaan uji regresi.

13.1.3. Kompetensi

Diakhir pembelajaran mahasiswa diharapkan mampu menjelaskan uji regresi linear sederhana

13.2. PENJELASAN

13.2.1. Uraian Isi

Analisis regresi digunakan untuk mengetahui variabel dependen/kriteria dapat diprediksikan melalui variabel independen atau prediktor secara individual. Dampak dari

penggunaan analisis regresi dapat digunakan untuk memutuskan apakah naik atau menurunkan keadaan variabel independen atau untuk meningkatkan keadaan variabel dependen dapat dilakukan dengan meningkatkan variabel independen atau sebaliknya. Korelasi & Regresi keduanya mempunyai hubungan yang sangat erat. Setiap regresi pasti ada korelasi, tetapi korelasi belum tentu dilanjutkan dengan regresi. Analisis regresi dilakukan bila hubungan dua variabel berupa hubungan kausal atau fungsional. Untuk menetapkan kedua variabel mempunyai hubungan kausal atau tidak, maka harus didasarkan pada teori atau konsep-konsep tentang dua variabel.

Regresi sederhana didasarkan pada hubungan fungsional ataupun kausal antara satu variabel independen dengan satu variabel dependen

Persamaan Umum Rumus Regresi Linier;

$$Y = a + bX$$

Dimana :

Y = Subyek dalam variabel dependen yg diprediksi

a = harga y bila x

b = angka arah/koefisien regresi bila b(+) maka naik, bila (-) maka terjadi penurunan

X = subyek pada variabel independen yang mempunyai nilai tertentu

Untuk menghitung nilai a dan b digunakan rumus :

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

$$a = \frac{\sum Y}{n} - b \cdot \frac{\sum X}{n}$$

A. Contoh Soal

Sebuah penelitian yang bertujuan untuk mengetahui hubungan antara pengeluaran konsumsi (Y) dengan pendapatan keluarga (X). Untuk itu diambil sampel acak sebanyak 10 keluarga untuk diwawancarai, dan dari hasil penelitian itu diperoleh data sebagai berikut:

Konsumsi (Y)	5	6	8	9	10	12	12	14	15	20
Pendapatan (X)	6	8	10	12	13	17	20	22	24	28

Berdasarkan data tersebut:

1. Dugalah persamaan regresi populaisnya
2. Berikan intervensi terhadap nilai b yang diperoleh
3. Dugalah rata-rata pengeluaran konsumsi bila pendapatan seorang keluarga 18

Penyelesaian :

1. Persamaan regresi populasi akan diduga dengan persamaan regresi sampelnya:

X_i	Y_i	X_i^2	Y_i^2	$X_i Y_i$
6	5	36	25	30
8	6	64	36	48
10	8	100	64	80
12	9	144	81	108
13	10	169	100	130
17	12	289	144	204
20	12	400	144	240
22	14	484	196	308
24	15	576	225	360
28	20	784	400	560
$\Sigma = 160$	$\Sigma = 111$	$\Sigma = 3046$	$\Sigma = 1415$	$\Sigma = 2068$

$$n=10$$

$$\bar{X} = \frac{\Sigma X_i}{n} = \frac{160}{10} = 16$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{111}{10} = 11,1$$

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

$$b = \frac{10 \cdot (2068) - (160)(111)}{10 \cdot (3046) - (160)^2} = 0,60$$

$$a = \bar{Y} - b\bar{X} = 11,1 - 0,60(16) = 1,50$$

jadi persamaan regresi sampel sebagai penduga regresi populasinya adalah:

$$y = a + bx = 1,5 + 0,6x$$

2. Nilai $b = 0,6$ memiliki arti bahwa bila pendapatan naik sebesar satu unit. Maka rata-rata pengeluaran konsumsi naik sebesar 0,6 unit.
3. Bila $X = 18$ maka $Y = \dots\dots$

$$Y = 1,5 + 0,6x = 1,5 + 0,6(18) = 12,3$$

Jadi bila pendapatan keluarga 18, maka rata-rata pengeluaran konsumsinya diharapkan sebesar 12,3

13.3. PENUTUP

Latihan

1. Sebuah sampel acak yang terdiri dari 6 pasangan data mengenai besarnya pendapatan dan konsumsi bulanan dari 6 karyawan perusahaan swasta yang bergerak di bidang pariwisata (dalam jutaan rupiah) adalah sebagai berikut:

Konsumsi (Y)	8	12	16	20	24	26
Pendapatan (X)	7	9	12	14	13	15

Berdasarkan data tersebut:

- a. Susunlah persamaan regresinya
- b. Berikan interpretasi terhadap nilai koefisien regresinya
- c. Taksirlah konsumsi seorang karyawan yang pendapatnya 23 juta rupiah

DAFTAR PUSTAKA

- Sudjana. 2005. Metoda statistika. Tarsito. Bandung
- Riwidikdo, H. 2012. Statistik kesehatan. Mitra cendikia press. Yogyakarta
- Budiarto, E. 2001. Biostatistika untuk kedokteran dan kesehatan masyarakat. EGC. Bandung
- Mason, R.D & Douglas A. Lind. 1996. Teknik statistik bisnis dan ekonomi. Erlangga. Jakarta
- Usman, H. 2000. Pengantar statistika. Bumi aksara. Bandung.
- Ating, S. 2006. Aplikasi statistika dalam penelitian. Pustaka setia. Bandung