



MODUL STATISTIK

Dr. PAULINA

MODUL STATISTIK

Dr. PAULINA



2015

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala limpahan Rahmat, Inayah, Taufik dan Hidayahnya sehingga kami dapat menyelesaikan penyusunan modul ini dalam bentuk maupun isinya yang sangat sederhana. Semoga modul ini dapat dipergunakan sebagai salah satu acuan, petunjuk maupun pedoman bagi mahasiswa dalam pengambilan matakuliah Statistik.

Harapan saya semoga modul ini membantu mahasiswa dalam mempelajari dan memahami matakuliah Statistik, dan pengalaman bagi para penyusun serta pembaca.

Modul ini kami akui masih banyak kekurangan karena pengalaman yang kami miliki sangat kurang. Oleh karena itu kami harapkan kepada para pembaca untuk memberikan masukan-masukan yang bersifat membangun untuk kesempurnaan modul ini, sehingga kami dapat memperbaiki bentuk maupun isi modul ini sehingga kedepannya dapat lebih baik.

Jakarta, November 2015

Penyusun

DAFTAR ISI

| | HAL | | |
|-----------------------------|----------------|--|------------|
| KATA PENGANTAR | i | | |
| DAFTAR ISI | ii | | |
| | | | |
| PERTEMUAN | CHAPTER | POKOK BAHASAN | HAL |
| 1 | 1 & 2 | Pengenalan Data dan Statistik | 3 |
| 2 | 3 | Statistik Deskriptif | 11 |
| 3 | 17 | Angka Indeks | 17 |
| 4 | 18 | Teknik Peramalan data | 21 |
| 5 | 14 | Konsep dasar Probabilitas | 26 |
| 6 | 5 | Distribusi probabilitas Diskrit dan Kontinue | 33 |
| 7 | 21 | Teknik Pengambilan Keputusan | 39 |
| 8 | 7 & 8 | Sampling dan Distribusi Sampling | 44 |
| 9 | 9 | Uji Hipotesis | 53 |
| 10 | 10 | Statistik Inference | 63 |
| 11 | 12 | Goodness of Fit Test & Test of Independence.... | 71 |
| 12 | 13 | Analisis Variance | 76 |
| 13 | 14 & 15 | Simple Kinier Regression & Multiple Regression | 83 |
| 14 | 19 | Non Parametric Methods | 94 |

SILABUS MATA KULIAH STATISTIK

1. Deskripsi dan Tujuan Mata Kuliah

1.1. Deskripsi Mata Kuliah

Melalui mata kuliah ini, mahasiswa akan mempelajari jenis-jenis data, bagaimana memperolehnya, dan bagaimana mengolah, menyajikan, hingga menginterpretasikan agar dapat dipergunakan untuk mengambil keputusan atas berbagai permasalahan yang ditemui sehari-hari.

Setelah memahami statistiK deskriptif, mahasiswa akan diberikan pemahaman tentang estimasi interval, uji hipotesa, statistic inference mengenai mean dan proporsi dengan dua populasi variance populasi. Selain itu, dipelajari juga tentang pengujian goodness of fit dan independence serta analisa variance. Materi tentang persamaan regresi baik simple maupun multiple terkait dengan penggunaan model seerta statistic non parametric. Secara keseluruhan, mata kuliah ini akan memberikan konsep dan kemampuan dalam menganalisis dan menginterpretasikan hasil pengolahan data statistik, dapat memprediksi parameter populasi serta menggunakan statistic sebagai salah satu alat analisis dalam pengambilan keputusan dari permasalahan ekonomi dan bisnis.

1.2. Capaian Pembelajaran

Setelah mengikuti perkuliahan Statistik, mahasiswa diharapkan mampu :

1. mengetahui ruang lingkup dan prinsip-prinsip dasar statistik
2. memahami konsep dan mampu menggunakan statistik dalam penyelesaian masalah
3. menggunakan metode statistik sebagai alat bantu dalam menganalisis dan menginterpretasikan hasil pengolahan data statistic, untuk keperluan penelitian dan pengambilan keputusan dalam berbagai permasalahan ekonomi dan bisnis.

2. Metodologi Pengajaran:

1. Perkuliahan akan dilakukan dalam 14 kali pertemuan, yang terbagi dalam 7 kali perkuliahan yang diikuti dengan Ujian Tengah Semester (UTS) dan kemudian dilanjutkan dengan 7 kali perkuliahan yang diakhiri dengan Ujian Akhir Semester (UAS).
2. Pokok bahasan untuk setiap pertemuan disusun sebagaimana jadwal terlampir. Mahasiswa diharapkan telah mempersiapkan dan membaca bahan bacaan yang

ditentukan sebelum mengikuti perkuliahan. Dengan demikian mahasiswa akan siap untuk mengikuti kuliah dengan lebih efektif.

3. Mahasiswa diharapkan berpartisipasi aktif dalam setiap tatap muka perkuliahan dengan membuat makalah dan bahan perkuliahan yang ditentukan sebagai bahan diskusi demi sukses dan tercapainya tujuan dari proses belajar mengajar ini.
4. Mahasiswa wajib mengerjakan seluruh tugas yang diberikan.
5. Kehadiran dalam perkuliahan minimum sejumlah 80 % dari total 14 kali pertemuan.
6. Tidak ada penundaan bagi tugas yang harus dikumpulkan.
7. Tidak ada kuliah tambahan di luar jadwal yang telah diberikan.

3. Bahan Bacaan

3.1. Bahan Bacaan Wajib

1. Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : *Statistics for Business and Economics*, 11th Edition, South-Western International Edition (AS)

3.2. Bahan Bacaan Pelengkap

1. Lind., Marchal., Wathen, 2008., : *Statistical Technique in Business and Economics With Global data Sets*, 13th Mc Graw Hill. (LD)

4. Evaluasi Hasil Belajar

| Unsur-unsur penilaian dan bobotnya | |
|------------------------------------|------|
| Tugas | 30 % |
| Ujian Tengah Semester (UTS) | 35 % |
| Ujian Akhir Semester (UAS) | 35 % |

5. Jadwal Pertemuan

| Pertemuan | Topik Bahasan | Bab |
|-----------|--|-------------------|
| 1 | Pengenalan data dan statistik | |
| | 1. Kegunaan statistic dalam bisnis dan ekonomi | AS.1.1 |
| | 2. Konsep dan jenis data | AS.1.2 |
| | 3. Sumber dan bagaimana memperoleh data, Skala Pengukuran | AS.1.3 |
| | 4. Pengenalan statistic deskriptif dan statistic inferensi | AS.1.4 AS.2.1. |
| | Statistik deskriptif dengan metode tabular dan grafik | AS.2.2 |

| | | |
|----------|--|--|
| | <ol style="list-style-type: none"> 1. Pengolahan data kualitatif (distribusi frekuensi, bar graph, pie chart) 2. Pengolahan data kuantitatif (distribusi frekuensi, histogram, distribusi kumulatif, ogive, scatter diagram) | AS.2.4 |
| 2 | <p>Statistik deskriptif dengan metode numerik</p> <ol style="list-style-type: none"> 1. Ukuran nilai sentral data 2. Ukuran variasi data 3. Ukuran variasi data (variance, standar deviasi, koefisien variasi) 4. Rata-rata tertimbang dan pengolahan data berkelompok | AS.3.1 AS.3.2 AS.3.6 |
| 3 | <p>Angka indeks</p> <ol style="list-style-type: none"> 1. Konsep harga relatif/indeks harga 2. Indeks harga gabungan, metode laspeyres dan Paasche 3. Penggunaan angka indeks pada perhitungan nilai uang riil | AS.17.1 AS.17.2 AS.17.5 |
| 4 | <p>Teknik Peramalan Data</p> <ol style="list-style-type: none"> 1. Data time series 2. Peramalan data dengan metode Moving Average dan Exponential Smoothing 3. Peramalan data dengan proyeksi trend | AS.18.1 AS.18.2 AS.18.3 |
| 5 | <p>Konsep dasar probabilitas</p> <ol style="list-style-type: none"> 1. Counting rule, kombinasi, permutasi 2. Probabilitas dari suatu event <p>Hubungan dasar probabilitas</p> <ol style="list-style-type: none"> 1. Komplemen dan hukum penjumlahan 2. Hukum perkalian dan probabilitas bersyarat 3. Teorema bayes | AS.4.1 AS.4.2 AS.4.3 AS.4.4 AS.4.5 |
| 6 | <p>Distribusi probabilitas diskrit</p> <ol style="list-style-type: none"> 1. Variabel random 2. Expected value dan varians <p>Distribusi probabilitas kontinue</p> <ol style="list-style-type: none"> 1. Distribusi probabilitas binomial, poisson, dan hipergeometrik 2. Distribusi probabilitas normal | AS.5.1 AS.5.2 AS.5.3 AS.5.4; AS.5.5; AS.5.6 |

| | | |
|------------------------------------|--|--|
| 7 | Teknik Pengambilan Keputusan 1. Formulasi masalah 2. Pengambilan keputusan dengan probabilitas Quis I | AS.21.1 AS.21.2 AS.21.3 AS.21.4 |
| UJIAN TENGAH SEMESTER (UTS) | | |
| 8 | Sampling dan Distribusi Sampling 1. Simple random sampling 2. Point estimation 3. Sampling distribution 4. Interval estimasi | AS.7 AS.8 |
| 9 | Uji Hipotesis 1. Membuat hipotesis nol dan alternatif 2. Kesalahan tipe I dan II 3. Uji rata-rata populasi : kasus sampel besar 4. Uji rata-rata populasi: kasus sampel kecil 5. Uji proporsi populasi 6. Uji hipotesis dan pengambilan keputusan | AS.9.1 AS.9.4 |
| 10 | Statistik Inference 1. Statistik inference tentang rata-rata dan proporsi dengan 2 populasi 2. Estimasi perbedaan rata-rata dua populasi: sampel independen 3. Uji hipotesis beda rata-rata antara dua populasi 4. Inference tentang variance populasi Quiz 2 | AS.10.1 AS.10.2 |
| 11 | Goodness of fit test & Test of independence 1. Uji goodness of fit test : populasi multinomial 2. Uji independensi 3. Uji goodness of fit test : distribusi normal | AS.12 |
| 12 | Analisis Variance 1. ANOVA : uji kesamaan k rata-rata populasi 2. Multiple comparison procedurs Quis 2 | AS.13 |
| 13 | Simple linier regression & Multiple regression 1. Metode least square 2. Koefisien determinasi 3. Uji signifikansi | AS.14 AS.15 |

| | | |
|-----------------------------------|---|-------|
| | <ul style="list-style-type: none"> 4. Penggunaan persamaan regresi untuk estimasi dan prediksi 5. Multiple coefficient of determinant 6. Uji signifikansi | |
| 14 | <p>Nonparametric methods</p> <ul style="list-style-type: none"> 1. Sign test with small sample case dan large sample case 2. Analisis residual 3. Wilcoxon signed rank test 4. Man-whitney –wilcoxon test 5. Kruskal walis test 6. Rank correlation | AS.19 |
| UJIAN AKHIR SEMESTER (UAS) | | |

Chapter 1 & 2 PENGENALAN DATA & STATISTIK

Capaian Pembelajaran Perkuliahan :

1. Memahami dan mampu menjelaskan statistik, data dan kegunaannya
2. Memahami dan mampu menjelaskan jenis-jenis data
3. Memahami cara memperoleh data dan skala pengukuran
4. Memahami dan mampu menjelaskan perbedaan antara statistik deskriptif dan inference

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

Applications in Business and Economics

1. Data
2. Data sources
3. Descriptive statistics
4. Statistical inference

Statistika adalah ilmu yang mengumpulkan, menata, menyajikan, menganalisis dan menginterpretasikan data menjadi informasi untuk membantu pengambilan keputusan yang efektif.

Statistik adalah suatu kumpulan angka yang tersusun lebih dari satu angka.

Data

Data adalah fakta dan angka yang dikumpulkan, dirangkum, dianalisis, dan ditafsirkan. Data yang dikumpulkan dalam studi tertentu yang disebut sebagai mengatur data.

Elemen, Variabel, dan Pengamatan

Elemen adalah entitas yang data dikumpulkan.

Variabel adalah karakteristik atau sifat dari masing-masing unit eksperimen atau sesuatu yang menarik dari elemen.

Observasi adalah pengukuran yang dikumpulkan dari elemen tertentu

Jumlah total nilai data dalam satu set data lengkap adalah jumlah elemen dikalikan dengan jumlah variabel.

Data, Data Sets, Elements, Variables, and Observations

The diagram illustrates a data set table with callouts identifying its components:

- Element Names:** Points to the 'Company' column.
- Observation:** Points to the rows of data.
- Variables:** Points to the columns 'Stock Exchange', 'Annual Sales(\$M)', and 'Earn/Share(\$)'.
- Data Set:** Points to the entire table.

| Company | Stock Exchange | Annual Sales(\$M) | Earn/Share(\$) |
|--------------|----------------|-------------------|----------------|
| Dataram | NQ | 73.10 | 0.86 |
| EnergySouth | N | 74.00 | 1.67 |
| Keystone | N | 365.70 | 0.86 |
| LandCare | NQ | 111.40 | 0.33 |
| Psychemedics | N | 17.60 | 0.13 |

Skala Pengukuran

- Skala nominal adalah angka yang diberikan kepada obyek mempunyai arti sebagai label saja, dan tidak menunjukkan tingkatan apa-apa.
- Skala ordinal adalah angka yang diberikan di mana angka-angka tersebut mengandung pengertian tingkatan.
- Skala interval adalah suatu skala pemberian angka pada obyek yang mempunyai sifat ukuran ordinal dan mempunyai jarak atau interval yang sama.
- Skala rasio adalah skala yang memiliki nilai nol dan rasio dua nilai yang memiliki arti.

Jenis-jenis Variabel

- Variabel kualitatif adalah data yang diperoleh dari sampel atau populasi berupa data kualitatif, data bukan berupa angka.
- Variabel kuantitatif adalah data yang diperoleh dari sampel atau populasi berupa data kuantitatif, data berupa angka.

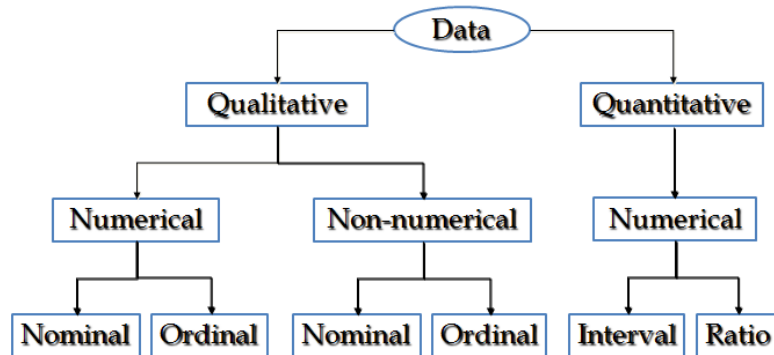
Data kualitatif

- Label atau nama yang digunakan untuk mengidentifikasi atribut masing-masing Element
- Sering disebut sebagai kategori Data
- Gunakan baik skala nominal atau ordinal pengukuran
- Bisa berupa angka atau nonnumeric
- Analisis statistik yang tepat agak terbatas

Data kuantitatif

- Menunjukkan berapa banyak atau berapa banyak: Data kuantitatif
- Sebuah. diskrit, jika mengukur berapa banyak
- terus menerus, jika mengukur berapa banyak
- Data kuantitatif selalu numerik.
- Operasi aritmatika biasa bermakna untuk data kuantitatif.

Scales of Measurement



Data Cross Sectional

Adalah data cross-sectional dikumpulkan pada saat yang sama atau sekitar titik yang sama dalam waktu.

Data Time Series

Adalah data yang dikumpulkan selama beberapa periode waktu tertentu

Sumber Data

- a. Data primer adalah data yang diperoleh secara langsung dari objek penelitian. Untuk data primer diperoleh dengan wawancara langsung, wawancara tidak langsung dan pengiriman kuisioner.
- b. Data sekunder adalah data yang diperoleh dari sumber lain yang sudah dipublikasikan. Untuk data sekunder dapat diperoleh dari sumber data seperti BPS, Bank Indonesia, majalah, jurnal, atau melihat dari website yang ada.

Statistika Deskriptif adalah metode statistika yang digunakan untuk menggambarkan atau mendeskripsikan data yang telah dikumpulkan menjadi sebuah informasi.

Statistika Inference adalah metode yang digunakan untuk mengetahui tentang sebuah populasi berdasarkan suatu sampel atau contoh dengan menganalisis dan menginterpretasikan data menjadi sebuah kesimpulan.

Pengertian Populasi dan Sampel.

Populasi adalah sebuah kumpulan dari semua kemungkinan orang-orang, benda-benda dan ukuran lain dari objek yang menjadi perhatian.

Sampel adalah suatu bagian dari populasi tertentu yang menjadi perhatian.

Statistik inferensi adalah proses menggunakan data yang diperoleh dari sampel untuk membuat estimasi dan uji hipotesis tentang karakteristik populasi

Sensus adalah mengumpulkan data populasi

Survei adalah contoh mengumpulkan data untuk sampel

Statistik deskriptif: Menampilkan Tabel dan Grafis

Meringkas Data untuk Variabel Kategori

1. Meringkas Data Kualitatif

a. Distribusi Frekuensi

Distribusi frekuensi adalah ringkasan tabel data yang menunjukkan frekuensi (atau nomor) dari item di setiap beberapa kelas agar tidak tumpang tindih. Tujuannya adalah untuk memberikan wawasan tentang data yang tidak dapat dengan cepat diperoleh dengan melihat hanya pada data asli.

b. Distribusi Frekuensi Relatif

Distribusi Frekuensi Relatif dari kelas adalah fraksi atau proporsi jumlah item data milik kelas. Sebuah distribusi frekuensi relatif adalah ringkasan tabel dari satu set data yang menunjukkan frekuensi relatif untuk masing-masing kelas.

c. Persentase Distribusi Frekuensi

Persentase distribusi frekuensi kelas adalah frekuensi relatif dikalikan dengan 100. Distribusi frekuensi persentase adalah ringkasan tabel dari satu set data yang menunjukkan frekuensi persen untuk masing-masing kelas.

d. Grafik Batang (Bar Graph)

Adalah suatu grafik yang menampilkan kelas-kelasnya pada sumbu horisontal dan frekuensi kelasnya pada sumbu vertikal. Frekuensi kelas ditampilkan secara proporsional sesuai dengan tinggi batangnya.

e. Diagram Pie

Adalah suatu diagram yang menunjukkan proporsi atau persentase yang diwakili tiap-tiap kelas atas jumlah total frekuensi.

- Pertama menggambar lingkaran; kemudian gunakan frekuensi relatif untuk membagi lingkaran menjadi sektor yang sesuai dengan frekuensi relatif untuk masing-masing kelas. –
- Karena ada 360 derajat dalam lingkaran, kelas dengan frekuensi relatif 0,25 akan mengkonsumsi $0,25 (360) = 90$ derajat lingkaran.

2. Meringkas Data Kuantitatif

a. Distribusi Frekuensi

Pengelompokkan data menjadi kelas-kelas yang tidak terikat satu sama lain yang menunjukkan jumlah pengamatan dalam tiap kelas.

Pedoman Memilih Jumlah Kelas:

1. Gunakan kelas antara 5 dan 20 kelas.
- 2) Data set dengan jumlah yang lebih besar dari elemen biasanya membutuhkan jumlah yang lebih besar dari kelas.
- 3) Data set yang lebih kecil biasanya membutuhkan kelas yang lebih sedikit.
- 4) Gunakan cukup kelas untuk menunjukkan variasi dalam data.

- 5) Jangan menggunakan begitu banyak kelas yang beberapa hanya berisi beberapa item data

Gunakan kelas dengan ukuran yang sama, menggunakan rumus :

$$\frac{\text{Largest Data Value} - \text{Smallest Data Value}}{\text{Number of Classes}}$$

b. Distribusi Frekuensi Relatif Kelas

Adalah frekuensi relatif adalah proporsi pengamatan dalam kelas. Dengan pengamatan sejumlah n,

$$\text{Relative frequency of class} = \frac{\text{frequency of the class}}{n}$$

c. Histogram

Adalah suatu grafik dimana kelas-kelas tercantum pada sumbu horozontal dan frekuensi kelasnya pada sumbu vertikal. Frekuensi kelas dilambangkan dengan tinggi batang dan batang-batangnya digambarkan saling menempel satu sama lain.

d. Distribusi Frekuensi Kumulatif

- 1) Distribusi frekuensi kumulatif menunjukkan jumlah item dengan nilai kurang dari atau sama dengan batas atas setiap kelas ..
- 2) Distribusi frekuensi relatif kumulatif menunjukkan proporsi item dengan nilai kurang dari atau sama dengan batas atas masing-masing kelas.
- 3) Disatribusi frekuensi kumulatif persentase menunjukkan persentase item dengan nilai kurang dari atau sama untuk batas atas masing-masing kelas.

e. Ogive

- 1) Ogive adalah grafik dari distribusi kumulatif.
- 2) Nilai data ditampilkan pada sumbu horisontal.
- 3) Pada sumbu vertikal menunjukkan:
 - frekuensi kumulatif, atau
 - frekuensi relatif kumulatif, atau
 - frekuensi kumulatif persen
- 4) Frekuensi (salah satu di atas) dari masing-masing kelas diplot sebagai titik.
- 5) Diplot poin dihubungkan dengan garis lurus.

Soal Latihan

1. Sebuah edisi terbaru Majalah Fortune melaporkan bahwa perusahaan berikut memiliki penjualan terendah per karyawan antara perusahaan Fortune 500.

| Company | Sales per Employee (In \$1,000s) | Sales Rank |
|---------|-------------------------------------|---------------|
|---------|-------------------------------------|---------------|

| | | |
|--------------------|-------|-----|
| Seagate Technology | 42.20 | 285 |
| SSMC | 42.19 | 414 |
| Russel | 41.99 | 480 |
| Maxxam | 40.88 | 485 |
| Dibrell Brothers | 22.56 | 470 |

- Berapa banyak elemen yang terdapat pada kumpulan data di atas?
- Berapa banyak variabel pada kumpulan data di atas?
- Berapa banyak pengamatan berada di kumpulan data di atas?
- Nama variabel dan menunjukkan data apakah mereka kualitatif atau kuantitatif.

2. Hasil nilai ujian statistik 40 mahasiswa STIE Indonesia Banking School adalah sebagai berikut :

| | | | |
|----|----|-----|----|
| 50 | 68 | 100 | 79 |
| 65 | 75 | 30 | 60 |
| 56 | 76 | 90 | 69 |
| 67 | 56 | 80 | 72 |
| 56 | 45 | 82 | 74 |
| 69 | 40 | 75 | 58 |
| 55 | 58 | 55 | 48 |
| 50 | 69 | 60 | 57 |
| 48 | 68 | 69 | 50 |
| 54 | 81 | 79 | 60 |

- Buatlah table frekuensi relative dan frekuensi persentasenya !
- Buatlah table frekuensi kumulatif relative dan persentase nya!
- Gambarlah grafik dari data tersebut!

3. Gaji bulanan dari 30 karyawan suatu perusahaan adalah sebagai berikut :

| Gaji bulanan (Ratusan ribu Rp.) | Banyaknya Karyawan |
|--|-------------------------------|
| (1) | (2) |
| 40 – 49 | 3 |
| 50 – 59 | 6 |
| 60 – 69 | 7 |
| 70 – 79 | 8 |
| 80 – 89 | 4 |
| 90 – 99 | 2 |

- Buatlah tabel frekuensi relative dan frekuensi persentase nya!
- Gambarlah grafik dari table diatas!

| |
|--|
| |
| |
| |

| Lembar Evaluasi | |
|---|--|
| Kriteria Penilaian: 4 : Apabila Skor latihan pertemuan 1, pengenalan data dan statistik benar 80 % > 80 3 : Apabila Skor latihan pertemuan 1, pengenalan data dan statistik benar 70 %: 70-79 2: Apabila Skor latihan pertemuan 1, pengenalan data dan statistik benar 50 %: 50-69 1: Apabila Skor latihan pertemuan 1, pengenalan data dan statistik benar 30 % : 31-59 0: Apabila Skor latihan pertemuan 1, pengenalan data dan statistik benar <30 | |
| Diperiksa oleh Asisten Dosen | Di <i>review</i> Oleh Dosen |
| Soal 1: nilai 80 Soal 2: nilai 80 Soal 3: nilai 80 Soal 4: nilai 80 (contoh) | Tingkat ketercapaian pembelajaran tergolong tuntas. Berdasarkan skor latihan rata-rata 80. Skor 4 |

Chapter 3 Statistik Deskriptif

Capaian Pembelajaran Perkuliahan :

1. Mampu menjelaskan dan menyelesaikan perhitungan ukuran nilai sentral dan variasi data
2. Mampu menjelaskan dan menyelesaikan perhitungan rata-rata tertimbang
3. Mampu menjelaskan dan menyelesaikan pengolahan data berkelompok

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Ukuran nilai sentral, variasi data
2. Perhitungan rata-rata tertimbang
3. Pengolahan data berkelompok

I. Ukuran Nilai Sentral, Variasi Data

a. Rata-Rata/Mean

Adalah rata-rata semua data sampel values. \bar{X} adalah titik estimator populasi yaitu μ
Rata-rata untuk sampel diberikan rumus

$$\bar{X} = \frac{\sum X_i}{n}$$

Rata-rata untuk Populasi diberikan rumus :

$$\mu = \frac{\sum X_i}{N}$$

b. Median

Adalah nilai di tengah-tengah atau titik tengah dari sejumlah nilai setelah nilai-nilai tersebut diurutkan dari yang terkecil hingga yang terbesar.

c. Mode/Modus

Adalah nilai pengamatan yang paling sering muncul atau frekuensi terbesar dapat terjadi pada dua atau lebih dari nilai modus yang berbeda. Jika data memiliki tepat dua mode, disebut data yang bimodal, dan jika data memiliki lebih dari dua mode, yang data multimodal.

d. Persentil

Sebuah persentil memberikan informasi tentang bagaimana data tersebar di interval dari nilai terkecil sampai yang terbesar. Persentil membagi bagian menjadi 100 bagian yang sama. Langkah-langkah dalam menghitung persentil

- 1) Susun data dalam urutan menaik
- 2) Hitung indeks i , posisi persentil ke p dengan rumus :

$$i = \left(\frac{p}{100}\right)n$$

- 3) Jika hasilnya bukan bilangan bulat, maka nilai ke P persentil adalah nilai dalam i posisi.
- 4) Jika hasilnya adalah bilangan bulat, persentil ke p adalah rata-rata nilai di posisi tersebut ditambah 1 ($i+1$).

e. Kuartil

Adalah membagi sekumpulan pengamatan menjadi empat bagian yang sama

- 1) Kuartil Pertama = 25 persen
- 2) Kuartil Kedua = 50 persen = Median
- 3) Kuartil Ketiga = 75 persen

b. Ukuran Variabilitas

Hal ini sering diinginkan untuk mempertimbangkan langkah-langkah variabilitas (dispersi), serta langkah-langkah dari lokasi.

- 1) Jarak (Range)

Kisaran data set adalah perbedaan antara nilai terbesar dan terkecil. Merupakan ukuran sederhana ukuran variabilitas.

- 2) Interkuartil Range

Merupakan perbedaan antara kuartil ketiga dengan kuartil pertama.

- 3) Varians

Varians adalah rata-rata hitung dari kuadrat deviasi rata-rata. Varians didasarkan pada perbedaan antara nilai pengamatan (x_i) dan mean (for \bar{X} sampel, μ untuk populasi) .

Varians Sampel :

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

Varians Populasi :

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

- 4) Standar Deviasi

Akar kuadrat dari variasi.

Standar deviasi dari data set adalah akar kuadrat positif dari variance. Standar deviasi diukur dalam satuan yang sama dengan data, sehingga lebih mudah diinterpretasikan dari varians.

$$\text{Varians untuk Sampel : } s = \sqrt{S^2}$$

$$\text{Varians untuk Populasi : } \sigma = \sqrt{\sigma^2}$$

5) Koefisien variasi

Koefisien variasi menunjukkan seberapa besar deviasi standar dalam kaitannya dengan mean.

$$\text{Koefisien Variasi untuk Sampel : } \left(\frac{s}{\bar{x}} \times 100\right) \%$$

$$\text{Koefisien Variasi untuk Populasi : } \left(\frac{\sigma}{\mu} \times 100\right) \%$$

II. Perhitungan Rata-Rata Tertimbang

1. Weigted Mean

Ketika rata-rata dihitung dengan memberikan bobot pada masing-masing nilai. Rata-rata tertimbang memperhitungkan dari nilai rata-rata (IPK), Bobot adalah jumlah kredit yang diterima untuk setiap grade. Rata-rata dengan bobot atau kepentingan dari setiap data berbeda. Besar dan kecilnya bobot tergantung pada alasan ekonomi dan teknisnya.

$$\bar{x} = \frac{\sum w_i x_i}{\sum w_i}$$

Data Berkelompok

Perhitungan rata-rata tertimbang dapat digunakan untuk memperoleh perkiraan dari mean, varians, dan standar deviasi untuk data yang dikelompokkan.

- 1) Untuk menghitung rata-rata tertimbang, kita memperlakukan titik tengah setiap kelas seolah-olah itu adalah mean dari semua item di kelas.
- 2) Kami menghitung rata-rata tertimbang dari titik tengah kelas menggunakan frekuensi kelas sebagai beban.
- 3) Demikian pula, dalam menghitung varians dan standar deviasi, frekuensi kelas yang digunakan sebagai beban.

Rumus Mean Data Berkelompok :

| | |
|-----------------|------------------------------------|
| Sample Data | $\bar{x} = \frac{\sum f_i M_i}{n}$ |
| Population Data | $\mu = \frac{\sum f_i M_i}{N}$ |

Rumus Varians Data Berkelompok :

sample data

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

population data

$$\sigma^2 = \frac{\sum f_i (M_i - \mu)^2}{N}$$

Soal Latihan

1. Berikut ini adalah data tinggi badan siswa-siswi SMA Bakti Mulya kelas XII :

Dari data disamping, tentukanlah :

| Tinggi (cm) | Frekuensi |
|-------------|-----------|
| 140 - 144 | 2 |
| 145 - 149 | 4 |
| 150 - 154 | 10 |
| 155 - 159 | 14 |
| 160 - 164 | 12 |
| 165 - 169 | 5 |
| 170 - 174 | 3 |

- Mean
- Median

2. Hitunglah nilai rata-rata dari nilai ujian matematika kelas 3 SMU berikut ini:

2; 4; 5; 6; 6; 7; 7; 7; 8; 9

- Mean
- Median
- Modus
- Quartil 1
- Quartil 3

3. Berikut adalah data nilai dengan bobot masing-masing

| Nilai (xi) | Bobot (wi) |
|------------|------------|
| 4.2 | 7 |
| 3 | 4 |
| 3.5 | 3 |
| 6 | 9 |

- Hitunglah weighted mean!
- Hitunglah mean dari sample disamping tanpa menggunakan bobot!

| |
|--|
| |
| |
| |
| |
| |
| |

| Lembar Evaluasi | |
|--|---|
| <p>Kriteria Penilaian:</p> <p>4 : Apabila Skor latihan pertemuan 2, Statistik Deskriptif 80 % > 80 3 : Apabila Skor latihan pertemuan 2, Statistik Deskriptif 70 %: 70-79 2: Apabila Skor latihan pertemuan 2, Statistik Deskriptif 50 %: 50-69 1: Apabila Skor latihan pertemuan 2, Statistik Deskriptif 30 % : 31-59 0: Apabila Skor latihan pertemuan 2, Statistik Deskriptif <30</p> | |
| Diperiksa oleh Asisten Dosen | Di <i>review</i> Oleh Dosen |
| Soal 1: nilai 80 Soal 2: nilai 80 Soal 3: nilai 80 Soal 4: nilai 80 (contoh) | Tingkat ketercapaian pembelajaran tergolong tuntas. Berdasarkan skor latihan rata-rata 80. Skor 4 |

Chapter 17

Angka Indek

Capaian Pembelajaran Perkuliahan :

1. Mampu menjelaskan dan menyelesaikan perhitungan konsep harga relatif
2. Mampu menjelaskan dan menyelesaikan perhitungan indeks harga gabungan, laspayres dan Pasche
3. Mampu menjelaskan dan menggunakan angka indeks pada perhitungan nilai uang riil

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Konsep harga relatif
2. Indeks harga gabungan, Laspayres dan Passche
3. Penggunaan angka indeks

Angka Indeks

Angka indeks yang menunjukkan perbandingan antara jumlah harga kelompok barang dan jasa pada periode tertentu dengan periode dasarnya

I. Harga Relatif

Harga relatif membantu dalam memahami dan menafsirkan perubahan kondisi ekonomi dan bisnis dari harga berlaku. Harga relatif menunjukkan bagaimana harga per unit untuk komoditas tertentu dibandingkan dengan harga periode dasar per unit untuk item yang sama. Harga relatif harga unit di setiap periode sebagai persentase dari harga unit pada periode periode dasar tertentu.

$$\text{Price relative in period } t = \frac{\text{Price in period } t}{\text{Base period price}} (100)$$

II. Indeks Harga Gabungan

Indeks harga agregat dikembangkan untuk tujuan khusus mengukur perubahan gabungan dari sekelompok barang. Indeks harga agregat tertimbang pada periode t , dinotasikan dengan rumus berikut :

$$I_t = \frac{\sum P_{it}}{\sum P_{i0}} (100)$$

Indeks Tertimbang

Indeks tertimbang memberikan bobot yang berbeda terhadap setiap komponen. Mengapa harus diberikan bobot yang berbeda? Karena pada dasarnya setiap barang dan jasa mempunyai tingkat utilitas (manfaat dan kepentingan) yang berbeda.

Dengan indeks agregat tertimbang setiap item dalam kelompok tersebut tertimbang menurut kepentingannya, yang biasanya adalah kuantitas yang dipergunakan. Dimana Q_i = kuantitas barang i , indeks harga agregat tertimbang pada periode t dinoatasikan dengan rumus :

Indeks Laspayres

Etienne Laspeyres mengembangkan metode ini pada abad 18 akhir untuk menentukan sebuah indeks tertimbang dengan menggunakan bobot sebagai penimbang adalah periode dasar.

$$I_t = \frac{\sum P_{it} Q_{i0}}{\sum P_{i0} Q_{i0}}$$

Indeks Passche

Menggunakan bobot tahun berjalan dan bukan tahun dasar sebagai bobot.

$$I_t = \frac{\sum P_{it} Q_i}{\sum P_{i0} Q_i} (100)$$

III. Penggunaan Angka Indeks

Angka indeks diperguanakn untuk berbagai macam keperluan baik bidang ekonomi maupun non ekonomi, terutama dengan melibatkan data yang besar. Penggunaan angka indeks di Indonesia antara lain :

1. Indeks Harga Konsumen
2. Indeks Harga Perdagangan Besar
3. Indeks Nilai Tukar Petani

| SEKTOR | Quantity | | Unit Price | |
|----------------|----------|--------|------------|-------|
| | 1994 | 2000 | 1994 | 2000 |
| Tempat Tinggal | 9,473 | 8,804 | 2.12 | 10.92 |
| Komersil | 5,416 | 6,015 | 1.97 | 11.32 |
| Industri | 21,287 | 17,832 | 0.79 | 5.13 |
| Transportasi | 15,293 | 20,262 | 2.32 | 6.16 |

4. Indeks Produktivitas

Soal Latihan

1. Berikut adalah data konsumsi dan pengeluaran energy

menurut sector di Kota Newton. Hitunglah Weighted Aggregate Price Index untuk pengeluaran energy pada tahun 2000 dengan **tahun dasar 1994!**

2. Berikut merupakan table harga dan kuantitas konsumsi barang A, B, dan C tahun 2014 dan 2015 adalah sbb :

| Macam barang | 2014 | | 2015 | |
|---------------|------------|----------------|------------|----------------|
| | Harga (Po) | Kuantitas (Qo) | Harga (Pt) | Kuantitas (Qt) |
| A | Rp 1000,- | 10 | Rp 1500,- | 5 |
| B | Rp 1500,- | 15 | Rp 1700,- | 10 |
| C | Rp 2000,- | 5 | Rp 2200,- | 4 |
| Jumlah | | | | |

Tentukan :

- Weighted aggregate price index dengan menggunakan laspeyres index
 - Weighted aggregate price index dengan menggunakan paasce index
3. Pada Pasar Kebayoran Lama terdaftar list perbandingan harga beberapa macam barang dari tahun 2003 dan tahun 2004.

| Macam Barang | Harga 2003 | Harga 2004 |
|--------------|-------------|-------------|
| A | Rp 200,00 | Rp 300,00 |
| B | Rp 300,00 | Rp 350,00 |
| C | Rp 500,00 | Rp 500,00 |
| D | Rp 100,00 | Rp 50,00 |
| E | Rp 200,00 | Rp 300,00 |
| Σ | Rp 1.300,00 | Rp 1.500,00 |

- Hitunglah Unweighted Aggregate Price Index dari daftar harga tersebut!
- Berapakah tingkat kenaikan atau

penurunannya

Lembar Jawaban

| |
|--|
| |
| |
| |
| |
| |
| |

| | |
|---|---|
| <p>4 : Apabila Skor latihan pertemuan 3, Angka Indeks 80 % > 80</p> <p>3 : Apabila Skor latihan pertemuan 3, Angka Indeks 70 %: 70-79</p> <p>2: Apabila Skor latihan pertemuan 3, Angka Indeks 50 %: 50-69</p> <p>1: Apabila Skor latihan pertemuan 3, Angka Indeks 30 % : 31-59</p> <p>0: Apabila Skor latihan pertemuan 3, Angka Indeks <30</p> | |
| Diperiksa oleh Asisten Dosen | <i>Di review Oleh Dosen</i> |
| <p>Soal 1: nilai 80</p> <p>Soal 2: nilai 80</p> <p>Soal 3: nilai 80</p> <p>Soal 4: nilai 80</p> <p>(contoh)</p> | <p>Tingkat ketercapaian pembelajaran tergolong tuntas. Berdasarkan skor latihan rata-rata 80.</p> <p>Skor 4</p> |

Chapter 18 Forecasting

Capaian Pembelajaran Perkuliahan :

1. Menguasai pengetahuan dan mampu menghitung peramalan dengan data time series
2. Menguasai pengetahuan dan mampu menghitung metode moving average dan exponential smoothing
3. Menguasai pengetahuan dan mampu menghitung peramalan dengan proyeksi trend

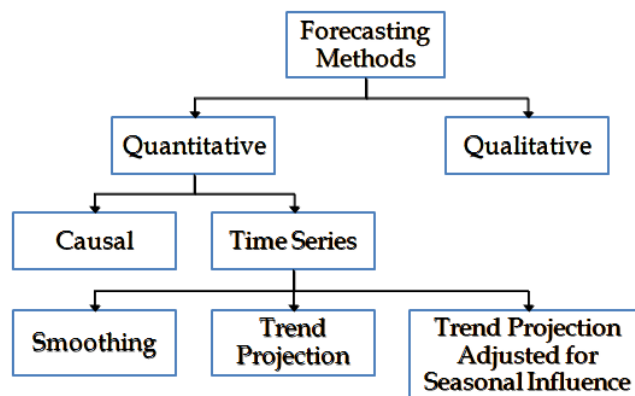
Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Data time series
2. Metode moving average dan exponential smoothing
3. Peramalan dengan proyeksi trend

Forecasting Methods



I. Pola Time Series

1. Tiga metode time series

a. Smoothing

Dalam kasus di mana data runtut waktu cukup stabil dan tidak memiliki tren yang signifikan, musiman, atau efek siklus, dapat menggunakan metode smoothing untuk rata-rata dari komponen yang tidak teratur dari seri waktu. Tiga metode smoothing umum adalah:

1. **Moving Averages**

Metode rata-rata bergerak terdiri dari perhitungan rata-rata nilai n . Rata-rata ini dipergunakan untuk peramalan nilai time series untuk periode berikutnya

$$\text{Moving Average} = \frac{\sum (\text{most recent } n \text{ data values})}{n}$$

2. Weighted Moving Averages

Untuk menggunakan metode ini pertama-tama kita harus memilih jumlah nilai data untuk dimasukkan dalam rata-rata. Berikutnya, kita harus memilih bobot untuk masing-masing nilai data. Berikutnya, kita harus memilih bobot untuk masing-masing nilai data. Pengamatan lebih baru biasanya diberikan bobot lebih dari pengamatan yang sebelumnya. Secara keseluruhan, bobot biasanya berjumlah 1.

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

b. Components of a Time Seri

Pola atau perilaku data dalam suatu kurun waktu memiliki beberapa komponen. Empat komponen kita akan mempelajari yang.



1. Component Trend/Trend Komponen

Perhitungan trend komponen untuk melihat pergeseran bertahap dari waktu ke waktu relatif lebih tinggi atau lebih rendah dalam jangka panjang. Trend biasanya merupakan hasil dari faktor jangka panjang seperti perubahan populasi, demografi, teknologi, atau preferensi konsumen.

2. Cyclical Component/Komponen Siklus

Setiap pola yang teratur dari urutan nilai-nilai di atas dan di bawah garis tren yang berlangsung lebih dari satu tahun dapat dikaitkan dengan komponen siklis. Biasanya, komponen ini adalah karena multi-gerakan siklis bertahun-tahun dalam perekonomian.

3. Seasonal Component/Komponen Musiman

Perhitungan komponen musiman untuk pola variabilitas teratur dalam jangka waktu tertentu, seperti variabilitas tahunan. Tidak selalu sesuai dengan musim tahunan (yaitu musim dingin, musim semi, musim panas, musim gugur). Dapat juga mingguan dan harian.

4. Irregular Component/Komponen Tidakberaturan

Komponen yang tidak teratur disebabkan oleh jangka pendek, faktor tak terduga dan tidak berulang yang mempengaruhi nilai dari waktu. Tidak bisa diprediksi.

c. Trend Projections

- 1) Jika kurun waktu menunjukkan trend linier, metode kuadrat dapat digunakan untuk menentukan garis tren (proyeksi) untuk prakiraan masa depan.
- 2) Kuadrat terkecil, juga digunakan dalam analisis regresi, menentukan perkiraan garis trend unik yang meminimalkan mean square error antara perkiraan garis trend dan nilai-nilai yang diamati sebenarnya untuk time series
- 3) Variabel bebas adalah periode waktu dan variabel dependen adalah nilai yang diamati sebenarnya dalam time series
- 4) Menggunakan metode kuadrat terkecil, rumus untuk proyeksi trend adalah

$$T_t = b_0 + b_1 t$$

$$b_1 = \frac{\sum tY_t - (\sum t \sum Y_t) / n}{\sum t^2 - (\sum t)^2 / n} \quad b_0 = \bar{Y} - b_1 \bar{t}$$

Soal Latihan

1. PT. Mira melakukan penjualan barang selama 8 tahun dengan jumlah unit sebagai berikut :

| YEAR | SALES |
|------|-----------|
| 2005 | Rp580,000 |
| 2006 | Rp620,000 |
| 2007 | Rp660,000 |
| 2008 | Rp700,000 |
| 2009 | Rp700,000 |
| 2010 | Rp720,000 |
| 2011 | Rp620,000 |
| 2012 | Rp600,000 |

| Waktu | Bulan | Penjualan |
|--------------|-----------|--------------|
| 1 | Januari | 1143 |
| 2 | Februari | 1037 |
| 3 | Maret | 857 |
| 4 | April | 757 |
| 5 | Mei | 948 |
| 6 | Juni | 660 |
| 7 | Juli | 683 |
| 8 | Agustus | 809 |
| 9 | September | 1078 |
| 10 | Oktober | 696 |
| 11 | November | 777 |
| 12 | Desembe | 672 |
| Total | | 10117 |

Dengan konstanta smoothing sebesar 0,2 , berapakah ramalan penjualan pada tahun 2013?

2. Penjualan Produk Foodbear pada tahun 2015 adalah sebagai berikut:

Tentukan peramalan penjualan pada bulan ke 13 dan 14 dengan Trend Projection!

Chapter 4

Konsep Dasar Probabilitas

Capaian Pembelajaran Perkuliahan :

1. Mampu menjelaskan dan menyelesaikan persoalan Kombinasi, permutasi
2. Mampu menjelaskan dan menyelesaikan persoalan Probabilitas suatu event
3. Mampu menjelaskan dan menyelesaikan Hukum penjumlahan, perkalian dan probabilitas bersyarat
4. Mampu menjelaskan dan menyelesaikan persoalan Teorema Bayes

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Kombinasi, permutasi
2. Probabilitas suatu event
3. Hukum penjumlahan, perkalian dan probabilitas bersyarat
4. Teorema Bayes

Definisi dan Manfaat

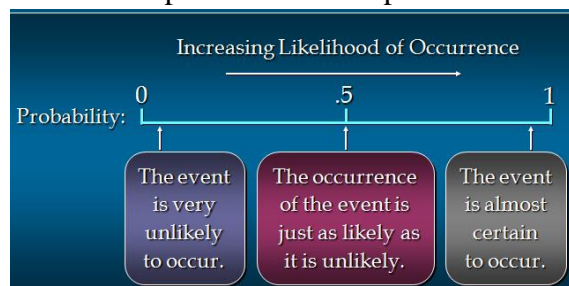
Definisi : Probabilitas adalah peluang suatu kejadian

Manfaat:

Manfaat mengetahui probabilitas adalah membantu pengambilan keputusan yang tepat, karena kehidupan di dunia tidak ada kepastian, dan informasi yang tidak sempurna.

Probabilitas:

Suatu ukuran tentang kemungkinan suatu peristiwa (*event*) akan terjadi di masa mendatang. Probabilitas dinyatakan antara 0 sampai 1 atau dalam persentase.



Percobaan:

Pengamatan terhadap beberapa aktivitas atau proses yang memungkinkan timbulnya paling sedikit dua peristiwa tanpa memperhatikan peristiwa mana yang akan terjadi.

Hasil (outcome):

Suatu hasil dari sebuah percobaan.

Peristiwa (event):

Kumpulan dari satu atau lebih hasil yang terjadi pada sebuah percobaan atau kegiatan.

Ruang Sampel (Sample space)

Sampel ruang untuk percobaan adalah himpunan hasil semua percobaan.

Aturan untuk menghitung : Langkah-langkah menghitung

- a. Jika percobaan terdiri dari urutan langkah-langkah k di mana ada hasil n1 mungkin untuk langkah pertama, Hasil n2 mungkin untuk langkah kedua, dan seterusnya, maka jumlah total hasil eksperimen adalah diberikan oleh (n1) (n2). . . (nk).
- b. Sebuah representasi grafis bermanfaat dari beberapa langkah eksperimen adalah dengan menggunakan diagram pohon.

Diagram pohon (tree diagram) adalah grafik yang berguna dalam menyusun perhitungan-perhitungan yang berlangsung dalam beberapa tahapan. Masing-masing bagian dalam pohon tersebut adalah satu tingkatan dari masalah. Cabang-cabang dari diagram pohon diberi bobot berdasarkan probabilitas.

I. Kombinasi, Permutasi

a. Kombinasi

Sebuah berguna aturan penghitungan kedua memungkinkan kita untuk menghitung jumlah hasil eksperimen saat n obyek untuk dipilih dari satu set N benda.

Jumlah Kombinasi N Objects Diambil n Sekaligus.

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!}$$

b. Permutasi

Aturan penghitungan ketiga berguna memungkinkan kita untuk menghitung jumlah hasil eksperimen saat n obyek untuk dipilih dari satu set N benda, di mana urutan seleksi adalah penting.

Jumlah Permutasi dari N Objects Diambil n Sekaligus

$$P_n^N = n! \binom{N}{n} = \frac{N!}{(N-n)!}$$

II. Probabilitas

a. Pendekatan Klasik

Setiap peristiwa mempunyai kesempatan yang sama untuk terjadi

Jika percobaan memiliki hasil n mungkin, metode ini akan menetapkan probabilitas $1/n$ untuk setiap hasil.

$$\text{Probabilitas suatu peristiwa} = \frac{\text{Jumlah kemungkinan hasil}}{\text{Jumlah total kemungkinan hasil}}$$

b. Pendekatan Relatif

Probabilitas suatu kejadian tidak dianggap sama, tergantung dari berapa banyak suatu kejadian terjadi. Probabilitas didapatkan dengan cara membagi frekuensi dengan total frekuensi.

$$\text{Probabilitas suatu peristiwa} = \frac{\text{Jumlah peristiwa yang terjadi}}{\text{Jumlah total percobaan}}$$

c. Pendekatan Subjektif

Probabilitas suatu kejadian didasarkan pada penilaian pribadi yang dinyatakan dalam suatu derajat kepercayaan.

- b. Ketika kondisi ekonomi dan perusahaan situasi berubah dengan cepat mungkin pantas untuk menetapkan probabilitas hanya berdasarkan data historis.
- c. Kita dapat menggunakan data yang tersedia serta kami pengalaman dan intuisi, tapi akhirnya probabilitas nilai harus mengungkapkan derajat kami keyakinan bahwa Hasil eksperimen akan terjadi.
- d. Perkiraan probabilitas terbaik sering diperoleh dengan menggabungkan perkiraan dari klasik atau relatif pendekatan frekuensi dengan estimasi subjektif.

d. Kejadian dan Probabilitas lainnya

- 1) Event adalah kumpulan titik sampel
- 2) Probabilitas setiap peristiwa adalah sama dengan jumlah dari probabilitas dari titik sampel dalam acara tersebut.
- 3) Jika kita dapat mengidentifikasi semua titik sampel dari bereksperimen dan menetapkan probabilitas untuk masing-masing, kita dapat menghitung probabilitas dari suatu peristiwa.

e. Beberapa Hubungan Dasar Probabilitas

Ada beberapa hubungan probabilitas dasar yang dapat digunakan untuk menghitung probabilitas dari suatu peristiwa tanpa pengetahuan tentang semua titik sampel probabilitas.

1. Complement of Event
 5. Komplemen dari kejadian A didefinisikan sebagai acara terdiri dari semua titik sampel yang tidak di A .
 6. Komplemen dari A dilambangkan dengan A^c .



2. Union of Two Events

7. Peristiwa A dan B adalah event yang mengandung semua titik sampel yang ada di A atau B atau keduanya.

8. Peristiwa A dan B dinotasikan dengan $A \cup B$



3. Intersection of Two Events

9. Perpotongan peristiwa A dan B adalah himpunan semua titik sampel yang baik A dan B.

10. Perpotongan peristiwa A dan B dinotasikan dengan $A \cap B$



III. Hukum Penjumlahan, Perkalian dan probabilitas Bersyarat

a. Addition Law

11. Hukum penjumlahan menyediakan cara untuk menghitung probabilitas kejadian A, atau B, atau A dan B terjadi.

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

12.

b. Mutually Exclusive Events

1) Dua peristiwa dikatakan saling eksklusif jika Peristiwa tidak memiliki titik sampel yang sama

2) Dua peristiwa yang saling eksklusif jika, ketika salah satu acara terjadi, yang lain tidak dapat terjadi.



3) Jika peristiwa A dan B saling eksklusif, $P(A \cap B) = 0$

4) Hukum penjumlahan saling eksklusif adalah:

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

c. Conditional Probability

- 1) Probabilitas suatu peristiwa mengingat bahwa acara lain telah terjadi disebut probabilitas bersyarat.
- 2) Probabilitas bersyarat dari A diberikan B dinotasikan oleh $P(A | B)$.
- 3) Sebuah peluang bersyarat dihitung sebagai berikut:

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

d. Multiplication Law

- 1) Hukum perkalian menyediakan cara untuk menghitung probabilitas persimpangan dua peristiwa.
- 2) Rumus hukum perkalian

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

e. Independent Events

- 1) Jika probabilitas dari kejadian A tidak berubah dengan Adanya acara B, kita akan mengatakan bahwa peristiwa A dan B adalah independen.
- 2) Dua peristiwa A dan B adalah independen jika:

$$P(A | B) = P(A) \quad \text{or} \quad P(B | A) = P(B)$$

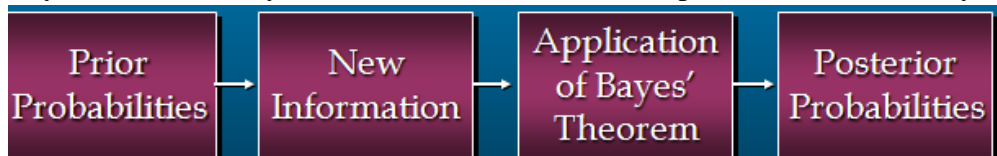
f. Multiplication Law for Independent Events

- 1) Hukum perkalian juga dapat digunakan sebagai tes untuk melihat jika dua peristiwa independen.
- 2) Dapat dirumuskan sebagai berikut :

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

IV. Bayes' Theorem

- 1) Seringkali kita mulai analisis probabilitas dengan awal atau probabilitas sebelumnya.
- 2) Kemudian, dari sampel, laporan khusus, atau kita memperoleh beberapa informasi tambahan.
- 3) Mengingat informasi ini, kita menghitung revisi atau probabilitas posterior.
- 4) Bayes 'teorema menyediakan sarana untuk merevisi probabilitas sebelumnya.



Soal Latihan

| | |
|------------------------------|--|
| Soal 4: nilai 80 (contoh) | |
|------------------------------|--|

Chapter 5 Distribusi Probabilitas Diskrit dan Kontinue

Capaian Pembelajaran Perkuliahan :

1. Mampu menjelaskan dan mengerjakan soal distribusi probabilitas diskrit
2. Mampu menjelaskan dan mengerjakan soal distribusi probabilitas kontinue
3. Mampu menjelaskan dan mengerjakan soal distribusi probabilitas normal

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Distribusi probabilitas diskrit
2. Distribusi probabilitas kontinue
3. Distribusi probabilitas normal

I. Distribusi Probabilitas Diskrit

- 1) Sebuah variabel acak adalah deskripsi numerik dari hasil percobaan.
- 2) Sebuah variabel acak diskrit mungkin menganggap baik jumlah terbatas nilai-nilai atau urutan tak terbatas nilai-nilai.
- 3) Sebuah variabel acak kontinu dapat berisi nilai numerik dalam selang waktu atau interval.
- 4) Distribusi probabilitas untuk variabel random menjelaskan bagaimana probabilitas didistribusikan lebih dari nilai variabel random
- 5) Distribusi probabilitas diskrit dapat digambarkan dengan tabel, grafik, atau persamaan.

a. Expected Value and Variance

Expected Value

Adalah Nilai yang diharapkan, atau ukuran lokasi pusat variabel acak

$$E(x) = \mu = \sum xf(x)$$

Variance

Varians adalah variabilitas dari nilai-nilai dari variabel random

$$\text{Var}(x) = \sigma^2 = \sum (x - \mu)^2 f(x)$$

Standar deviasi, σ , didefinisikan sebagai nilai positif akar kuadrat dari varians

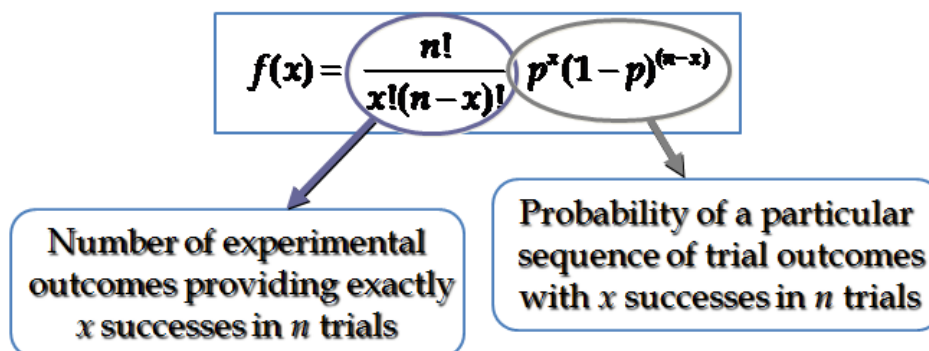
II. Distribusi Probabilitas Kontinue

1. Distribusi Binomial

Distribusi diskrit yang sering terjadi. Salah satu ciri distribusi binomial adalah hanya memiliki dua hasil yang mungkin terjadi dalam sebuah percobaan dari satu eksperimen.

- 1) Hasil setiap percobaan dari sebuah eksperimen dikategorikan menjadi salah satu dari dua kategori yang tidak terikat satu sama lain, yaitu : Sukses atau gagal
- 2) Variabel acak/random menghitung jumlah sukses yang muncul pada sejumlah percobaan
- 3) Probabilitas untuk sukses dan gagal selalu sama untuk setiap percobaan
- 4) Percobaan-percobaan saling bebas, yang artinya hasil dari satu percobaan tidak mempengaruhi hasil percobaan yang lain

■ Binomial Probability Function



Expected Value

$$E(x) = \mu = np$$

Variance

$$\text{Var}(x) = \sigma^2 = np(1 - p)$$

Standard Deviation

$$\sigma = \sqrt{np(1 - p)}$$

2. Distribusi Poisson

Menjelaskan berapa kali sebuah kejadian terjadi selama interval tertentu. Interval tersebut dapat berupa jarak, waktu, luas, atau volume

- 1) Variabel acaknya adalah berapa banyak sebuah kejadian terjadi selama interval yang ditentukan
- 2) Probabilitas kejadian tersebut proporsional dengan ukuran interval
- 3) Tidak ada pengulangan interval dan interval-intervalnya saling bebas

Poisson Probability Function

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

where:

$f(x)$ = probability of x occurrences in an interval

μ = mean number of occurrences in an interval

$e = 2.71828$

A property of the Poisson distribution is that the mean and variance are equal.

$$\mu = \sigma^2$$

3. Distribusi Hipergeometrik

Distribusi hipergeometrik berkaitan erat untuk distribusi binomial

- 1) Sebuah hasil untuk setiap percobaan dari sebuah eksperimen diklasifikasikan menjadi satu dari dua kategori tidak terikat satu sama lain : sukses atau gagal
- 2) Variabel acaknya adalah jumlah sukses pada sejumlah percobaan
- 3) Percobaan-percobaan itu tidak saling bebas
- 4) Asumsi sebuah populasi yang dapat dihitung tanpa pengembalian dan $n/N > 0.05$. jadi probabilitas sukses berubah pada setiap percobaan.

Hypergeometric Probability Function

$$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}} \quad \text{for } 0 \leq x \leq r$$

where: $f(x)$ = probability of x successes in n trials
 n = number of trials
 N = number of elements in the population
 r = number of elements in the population labeled success

Mean

$$E(x) = \mu = n \left(\frac{r}{N} \right)$$

Variance

$$Var(x) = \sigma^2 = n \left(\frac{r}{N} \right) \left(1 - \frac{r}{N} \right) \left(\frac{N-n}{N-1} \right)$$

4. Distribusi Normal

Ciri-ciri distribusi normal :

- 1) Kurva berbentuk lonceng (bell-shaped) dan memiliki satu puncak pada bagian tengah distribusi. Rata-rata, median dan modusnya sama dan terletak di pusat distribusi. Luas total di bawah kurva adalah 1. Setengah dari luas di bawah kurva normal ada di sebelah kanan dari titik pusatnya dan setengah yang lain ada di sebelah kiri.
- 2) Kurvanya simetris dengan subu di sekitra nilai rata-rata
- 3) Distribusinya asimtotik (menurun secara halus pada kedua arah dari bagian tengah)
- 4) Lokasi sebuah distribusi normal ditentukan oleh rata-rata μ .

Normal Probability Density Function

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

where:

μ = mean

σ = standard deviation

π = 3.14159

e = 2.71828

- 5) Variabel acak dengan rata-rata 0 dan standar deviasi 1, disebut sebagai distribusi probabilitas normal standar

| | |
|------------------------------|--|
| Soal 4: nilai 80 (contoh) | |
|------------------------------|--|

Chapter 21 Teknik Pengambilan Keputusan

Capaian Pembelajaran Perkuliahan :

1. Mampu menjelaskan dan melakukan formulasi masalah
2. Mampu menjelaskan dan mengambil keputusan dengan teknik probabilitas

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Formulasi masalah
2. Pengambilan keputusan dengan teknik probabilitas

I. Formulasi Masalah

a. Identifikasi hal-hal berikut :

- 1) Tentukan alternatif keputusan
- 2) Tentukan pernyataan kondisi alamiah (kejadian yang tidak pasti)
- 3) Tentukan hasil (payoff) yang dikombinasikan dengan : alternatif keputusan dan kondisi alamiah
- 4) Payoff dapat berupa : profit, biaya, waktu, jarak dan lain-lain

b. Pohon Diagram (Diagram Tree)

- 1) Sebuah pohon keputusan memberikan representasi grafis yang menunjukkan sifat berurutan dari proses pengambilan keputusan
- 2) Setiap pohon keputusan memiliki dua jenis node:
 - node putaran sesuai dengan kondisi alamiah

- node persegi sesuai dengan alternatif keputusan
- 3) Cabang-cabang meninggalkan setiap node putaran mewakili bagian yang berbeda dari kondisi alamiah sedangkan cabang meninggalkan setiap node persegi mewakili alternatif keputusan yang berbeda
 - 4) Pada akhir setiap dahan pohon adalah hasil yang diperoleh dari serangkaian cabang yang membentuk tungkai itu.

II. Pengambilan Keputusan Dengan Probabilitas

1. Pendekatan Probabilitas

- 1) Setelah menetapkan alternatif keputusan dan kondisi alamiah untuk setiap kesempatan, dan menentukan probabilitas untuk kondisi alamiah
- 2) Probabilitas dapat digunakan Metode klasik, metode frekuensi relatif, atau metode subjektif dari menugaskan
- 3) Karena satu dan hanya satu dari N-kondisi alamiah dapat terjadi, probabilitas harus memenuhi dua kondisi

$$P(s_j) \geq 0 \quad \text{for all states of nature}$$

$$\sum_{j=1}^N P(s_j) = P(s_1) + P(s_2) + \dots + P(s_N) = 1$$

- 4) Kemudian kita menggunakan pendekatan nilai yang diharapkan untuk mengidentifikasi yang terbaik atau alternatif keputusan yang direkomendasikan
- 5) Nilai yang diharapkan dari setiap alternatif keputusan dihitung
- 6) Keputusan alternatif yang menghasilkan nilai terbaik diharapkan akan dipilih.

2. Pendektanan Expected Value

- 1) Nilai yang diharapkan dari alternatif keputusan adalah jumlah hasil tertimbang untuk alternatif keputusan.
- 2) Nilai yang diharapkan (EV) dari alternatif keputusan di didefinisikan sebagai

The expected value of a decision alternative is the sum of weighted payoffs for the decision alternative.

The expected value (EV) of decision alternative d_i is defined as

$$\rightarrow EV(d_i) = \sum_{j=1}^N P(s_j)V_{ij}$$

where: N = the number of states of nature

$P(s_j)$ = the probability of state of nature s_j

V_{ij} = the payoff corresponding to decision alternative d_i and state of nature s_j

Langkah-langkah pendekatan Expected Value

- 1) Menghitung nilai yang diharapkan untuk setiap keputusan

- 2) Pohon keputusan pada slide berikutnya dapat membantu dalam perhitungan ini.
- 3) Berikut d1, d2, d3 mewakili alternatif keputusan model A, B, dan C.
- 4) Dan s1, s2, s3 mewakili kondisi alamiah

3. Expected Value of Perfect Information

- 1) Sering, informasi yang tersedia yang dapat meningkatkan perkiraan probabilitas untuk kondisi alamiah.
- 2) Nilai yang diharapkan dari informasi sempurna (EVPI) adalah peningkatan keuntungan yang diharapkan yang akan terjadi jika salah satu tahu diketahui dengan pasti dari keadaan alamiah yang akan terjadi.
- 3) EVPI memberikan informasi batas atas nilai yang diharapkan dari sampel atau survei
- 4) Nilai yang diharapkan dari informasi yang sempurna didefinisikan

$$EVPI = |EVwPI - EVwoPI|$$

where:

EVPI = expected value of perfect information

EVwPI = expected value *with* perfect information about the states of nature

EVwoPI = expected value *without* perfect information about the states of nature

Langkah-langkah menghitung EVPI

- 1) Menentukan laba yang optimal sesuai untuk setiap keadaan alamiah.
- 2) Menghitung nilai pengembalian optimal yang diharapkan
- 3) Kurangi EV keputusan optimal dari jumlah yang ditentukan pada langkah (2).
- 4) Menghitung nilai yang diharapkan untuk hasil optimal untuk setiap keadaan alamiah dan kurangi EV keputusan yang optimal

Soal Latihan

1. Asumsikan Anda memiliki sejumlah uang yang tersedia yang Anda ingin diinvestasikan di salah satu dari tiga rencana investasi yang tersedia: saham, obligasi, atau pasar uang. Hasil dari setiap rencana di bawah dua kondisi ekonomi yang mungkin adalah sebagai berikut. Probabilitas terjadinya kondisi ekonomi adalah 0,28.

| | Economic Condition I | Economic Condition II |
|---------------------|-----------------------------|------------------------------|
| Stocks | 1000 | 3000 |
| Bonds | 2500 | 2000 |
| Money Market | 1800 | 4000 |

- a. Hitunglah nilai yang diharapkan dari tiga pilihan investasi
- b. Hitunglah nilai yang diharapkan berdasarkan informasi yang sempurna

| | |
|------------------------------|--|
| Soal 4: nilai 80 (contoh) | |
|------------------------------|--|

Chapter 7 dan 8 Sampling dan Distribusi Sampling

Capaian Pembelajaran Perkuliahan :

1. Mampu menjelaskan dan melakukan perhitungan estimasi titik
2. Mampu menjelaskan dan melakukan perhitungan distribusi sampling
3. Mampu menjelaskan dan melakukan perhitungan interval estimate

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Simple random sampling
2. Point estimation
3. Sampling distribution
4. Interval estimate

I. Simple Randomn Sampling

a. Populasi Terbatas

- 1) Populasi terbatas sering didefinisikan oleh daftar seperti: daftar keanggotaan organisasi keanggotaan, Nomor rekening kartu kredit, Nomor persediaan produk
- 2) Sebuah sampel acak sederhana dari ukuran n dari populasi terbatas populasi berukuran N adalah sampel yang dipilih sedemikian rupa sehingga setiap kemungkinan sampel berukuran n memiliki probabilitas sama untuk terpilih.
- 3) Menggantikan setiap elemen sampel sebelum memilih elemen berikutnya disebut sampling dengan pengembalian

- 4) Sampling tanpa pengembalian jarang dipergunakan
- 5) Dalam proyek-proyek besar sampling, komputer yang dihasilkan nomor acak sering digunakan untuk mengotomatisasi Proses pemilihan sampel.

b. Populasi Tidak Terbatas

- 1) Populasi tak terbatas sering didefinisikan oleh proses yang berkelanjutan dimana elemen populasi terdiri dari item yang dihasilkan seolah-olah proses akan beroperasi tanpa batas.
- 2) Sebuah sampel acak sederhana dari populasi tak terbatas adalah sampel yang dipilih bahwa kondisi berikut seperti puas. Setiap elemen yang dipilih berasal dari sama populasi. Setiap elemen yang dipilih secara independen
- 3) Dalam kasus populasi yang tak terbatas, tidak mungkin untuk mendapatkan daftar semua elemen dalam populasi.
- 4) Prosedur seleksi nomor acak tidak dapat digunakan untuk populasi tak terbatas

II. Point Estimation

- 1) Dalam estimasi titik kita menggunakan data dari sampel untuk menghitung nilai dari statistik sampel yang berfungsi sebagai perkiraan parameter populasi.
- 2) Kita sebut \bar{X} sebagai titik estimator populasi berarti μ .
- 3) s adalah titik estimator standar deviasi popuasi σ
- 4) p adalah titik estimator proporsi populasi p

III. Sampling Distribution

Sampling Error

- 1) Nilai absolut dari perbedaan antara berisi titik estimasi dan sesuai parameter populasi disebut sampling error.

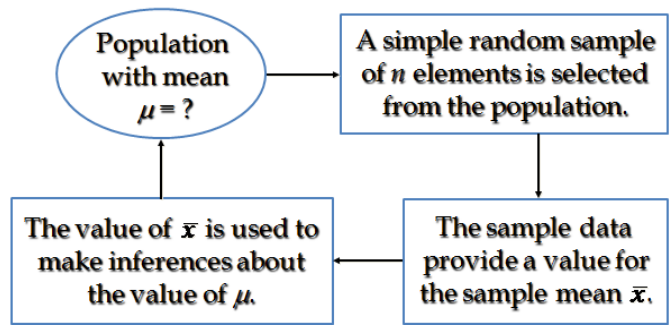
The sampling errors are:

| |
|---|
| $\bar{x} - \mu$ for sample mean |
|---|

| |
|--|
| $s - \sigma$ for sample standard deviation |
|--|

| |
|---|
| $\bar{p} - p$ for sample proportion |
|---|

■ Process of Statistical Inference



Expected Value of \bar{x}

$$E(\bar{x}) = \mu$$

where:

μ = the population mean

Standard Deviation of \bar{x}

Finite Population

$$\sigma_{\bar{x}} = \left(\frac{\sigma}{\sqrt{n}}\right) \sqrt{\frac{N-n}{N-1}}$$

Infinite Population

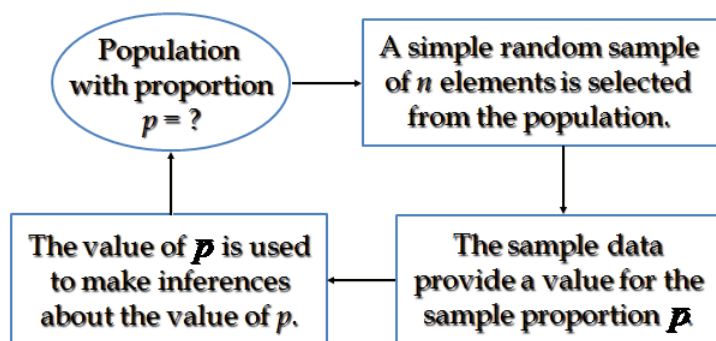
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

- A finite population is treated as being infinite if $n/N \leq .05$.
- $\sqrt{(N-n)/(N-1)}$ is the finite correction factor.
- $\sigma_{\bar{x}}$ is referred to as the standard error of the mean.

- Ketika populasi memiliki distribusi normal, distribusi sampling dari biasanya didistribusikan untuk setiap ukuran sampel.
- Dalam sebagian besar aplikasi, distribusi sampling dari dapat didekati dengan distribusi normal setiap kali sampel adalah ukuran 30 atau lebih.
- Dalam kasus di mana populasi sangat miring atau outlier yang hadir, sampel ukuran 50 mungkin diperlukan.

Sampling Distribution of \bar{p}

■ Making Inferences about a Population Proportion



Sampling Distribution of \hat{p}

Standard Deviation of \hat{p}

Finite Population

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \sqrt{\frac{N-n}{N-1}}$$

Infinite Population

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

$\sigma_{\hat{p}}$ is referred to as the standard error of the proportion.

Properties of Point Estimators

Sebelum menggunakan sampel statistik sebagai titik estimator, statistik memeriksa untuk melihat apakah sampel statistik memiliki sifat sebagai berikut terkait dengan titik estimator yang baik. Tidak Bias, Efisien dan Konsisten

Tidak Bias

Jika nilai yang diharapkan dari statistik sampel adalah sama dengan parameter populasi yang diperkirakan, statistik sampel dikatakan estimator berisi parameter populasi

Efisien

Diberi pilihan dua estimator berisi parameter populasi yang sama, kita akan lebih suka menggunakan jalur estimator dengan standar deviasi yang lebih kecil, karena cenderung memberikan perkiraan lebih dekat dengan parameter populasi.

Titik estimator dengan standar deviasi lebih kecil dikatakan memiliki efisiensi relatif lebih besar dari yang lain.

Konsisten

Sebuah titik estimator konsisten jika nilai-nilai titik estimator cenderung menjadi lebih dekat dengan parameter populasi sebagai ukuran sampel menjadi lebih besar

IV. Interval Estimate

a. Population Mean: σ Known

Margin of Error and the Interval Estimate

- 1) Titik estimator tidak dapat diharapkan untuk memberikan nilai yang tepat dari parameter, populasi
- 2) Perkiraan interval dapat dihitung dengan menambahkan dan mengurangi margin of error estimasi titik.

$$\text{Point Estimate } \pm \text{ Margin of Error}$$

- 3) Tujuan dari perkiraan interval untuk memberikan informasi tentang seberapa dekat estimasi titik adalah untuk nilai parameter

The general form of an interval estimate of a population mean is

$$\bar{x} \pm \text{Margin of Error}$$

b. Interval Estimation of a Population Mean: σ unKnown

- 1) Dalam rangka mengembangkan perkiraan interval populasi berarti, margin of error harus dihitung dengan menggunakan salah satu dari: standar deviasi populasi σ , atau standar deviasi sampel s .
- 2) σ jarang diketahui persis, tetapi sering perkiraan yang baik dapat diperoleh berdasarkan data historis atau informasi lainnya

Interval Estimate of μ

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

- where:
- \bar{x} is the sample mean
 - $1 - \alpha$ is the confidence coefficient
 - $z_{\alpha/2}$ is the z value providing an area of $\alpha/2$ in the upper tail of the standard normal probability distribution
 - σ is the population standard deviation
 - n is the sample size

c. Interval Estimation of a Population Mean: σ Unknown

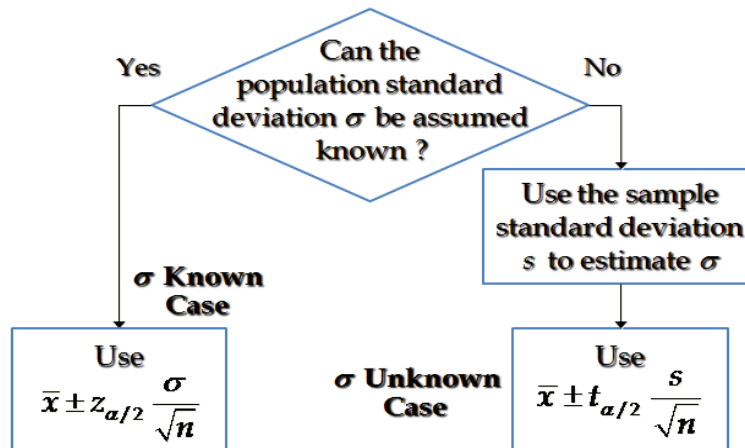
- 1) Jika perkiraan populasi standar deviasi σ tidak dapat dikembangkan sebelum sampling, kita menggunakan sampel standar deviasi s untuk memperkirakan σ
- 2) Dalam hal ini, bentuk estimasi interval berdasarkan distribusi t .
- 3) Sebuah distribusi t tertentu tergantung pada parameter yang dikenal sebagai derajat kebebasan.
- 4) Derajat kebebasan mengacu pada jumlah potongan independen informasi yang masuk ke perhitungan s .

Interval Estimate

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

- where:
- $1 - \alpha$ = the confidence coefficient
 - $t_{\alpha/2}$ = the t value providing an area of $\alpha/2$ in the upper tail of a t distribution with $n - 1$ degrees of freedom
 - s = the sample standard deviation

Summary of Interval Estimation Procedures for a Population Mean



Sample Size for an Interval Estimate of a Population Mean

Margin of Error

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Necessary Sample Size

$$n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2}$$

Interval Estimation of a Population Proportion

The general form of an interval estimate of a population proportion is

$$\bar{p} \pm \text{Margin of Error}$$

- 1) Distribusi sampling memainkan peran kunci dalam menghitung margin of error estimasi interval ini.
- 2) Distribusi sampling dapat didekati dengan distribusi normal setiap kali $np > 5$ dan $n(1 - p) > 5$.

Interval Estimation of a Population Proportion

Interval Estimate

$$\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

where: $1 - \alpha$ is the confidence coefficient

$z_{\alpha/2}$ is the z value providing an area of $\alpha/2$ in the upper tail of the standard normal probability distribution

\bar{p} is the sample proportion

Sample Size for an Interval Estimate of a Population Proportion

Margin of Error

$$E = z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Solving for the necessary sample size, we get

$$n = \frac{(z_{\alpha/2})^2 \bar{p}(1-\bar{p})}{E^2}$$

However, \bar{p} will not be known until after we have selected the sample. We will use the planning value p^* for \bar{p}

Sample Size for an Interval Estimate of a Population Proportion

Necessary Sample Size

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2}$$

The planning value p^* can be chosen by:

1. Using the sample proportion from a previous sample of the same or similar units, or
2. Selecting a preliminary sample and using the sample proportion from this sample.

Soal Latihan

1. ABC Company memproduksi 'Remote Control' dengan menggunakan dua baterai. Rata-rata umur baterai yang digunakan di produk ini adalah 35 jam. Distribusi umur baterai

Chapter 9

Uji Hipotesis

Capaian Pembelajaran Perkuliahan :

1. Mampu mengerjakan dan menganalisa hipotesis
2. Mampu mengerjakan dan menganalisa kesalahan Tipe I dan II
3. Mampu mengerjakan dan menganalisa pengambilan keputusan

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Membuat hipotesis nol dan hipotesis alternatif
2. Kesalahan Tipe I dan Tipe II
3. Uji rata-rata populasi : Kasus sampel besar
4. Uji rata-rata populasi : Kasus sampel kecil
5. Uji proporsi populasi
6. Uji hipotesis dan pengambilan keputusan

Pendahuluan :

Hipotesis adalah sebuah pernyataan tentang sebuah parameter populasi

Pengujian hipotesis adalah suatu prosedur berdasarkan bukti sampel dan teori probabilitas untuk menentukan apakah suatu hipotesis benar atau tidak.

I. Membuat hipotesis nol dan hipotesis alternatif

a. Membuat Hipotesis Nol dan Hipotesis Alternatif

- 1) Pengujian hipotesis dapat digunakan untuk menentukan apakah sebuah pernyataan tentang nilai parameter populasi harus atau tidak harus ditolak

- 2) Hipotesis nol, dinotasikan dengan H_0 , adalah asumsi tentatif tentang parameter populasi.
- 3) Hipotesis alternatif, dilambangkan dengan H_a , adalah kebalikan dari apa yang dinyatakan dalam hipotesis nol.

b. Pengujian Hipotesis Penelitian

- 1) Hipotesis alternatif adalah uji terhadap Hipotesis Penelitian
- 2) Kesimpulan bahwa hipotesis penelitian ini benar berasal dari data sampel yang bertentangan dengan hipotesis nol

c. Pengujian validitas

- 1) Kondisi dari keraguan dinyatakan sebagai hipotesis nol.
- 2) Kesimpulan tidak bahwa klaim tersebut palsu berasal dari data sampel yang bertentangan dengan hipotesis nol

d. Pengujian Dalam Pengambilan Keputusan

- 1) Pembuat keputusan mungkin harus memilih antara dua hal, salah satu yang terkait dengan hipotesis nol dan lainnya yang terkait dengan hipotesis alternatif
- 2) Secara umum, uji hipotesis tentang nilai rata-rata populasi berarti $\square\square$ harus mengambil salah satu dari tiga bentuk berikut (di mana μ_0 adalah nilai hipotesis dari rata-rata populasi)

| | | |
|--|--|--|
| $H_0: \mu \geq \mu_0$ $H_a: \mu < \mu_0$ | $H_0: \mu \leq \mu_0$ $H_a: \mu > \mu_0$ | $H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$ |
| One-tailed (lower-tail) | One-tailed (upper-tail) | Two-tailed |

II. Kesalahan Tipe I dan Tipe II

a. Kesalahan Tipe I

- 1) Penolakan hipotesis nol, ketika hipotesis tersebut benar
- 2) Probabilitas membuat kesalahan Tipe I ketika hipotesis nol benar sebagai kesetaraan disebut tingkat signifikansi.

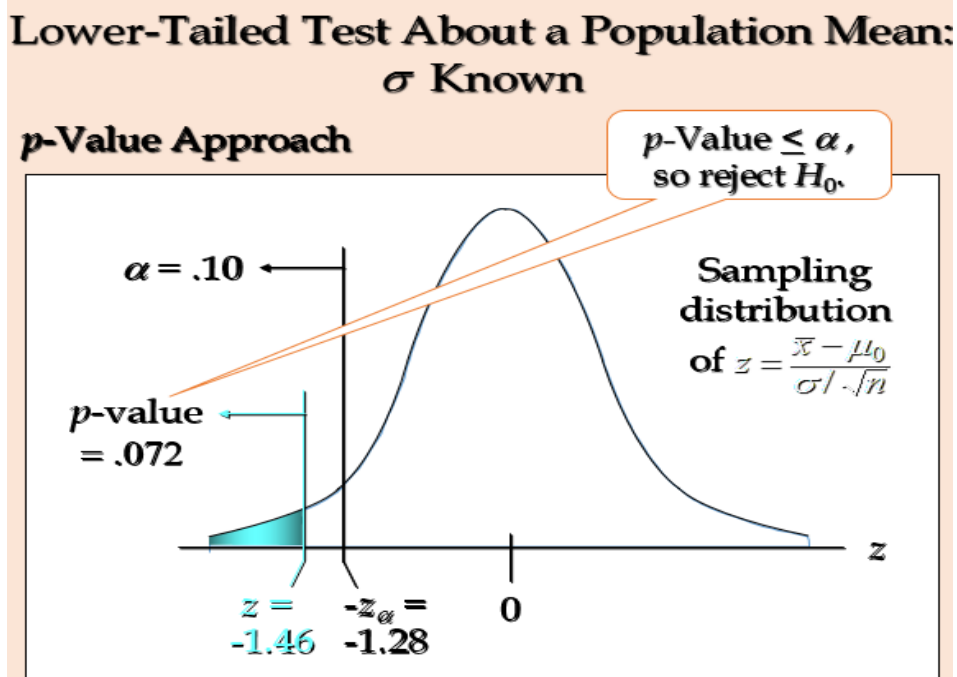
b. Kesalahan Tipe II

- 1) Penerimaan hipotesis nol, ketika hipotesis tersebut salah
- 2) Uji statistik menghindari risiko membuat kesalahan Tipe II dengan menggunakan "tidak menolak H_0 " dan tidak "menerima H_0 "

| Conclusion | Population Condition | |
|---|---------------------------------|-------------------------------|
| | H_0 True ($\mu \leq 12$) | H_0 False ($\mu > 12$) |
| Accept H_0 (Conclude $\mu \leq 12$) | Correct Decision | Type II Error |
| Reject H_0 (Conclude $\mu > 12$) | Type I Error | Correct Decision |

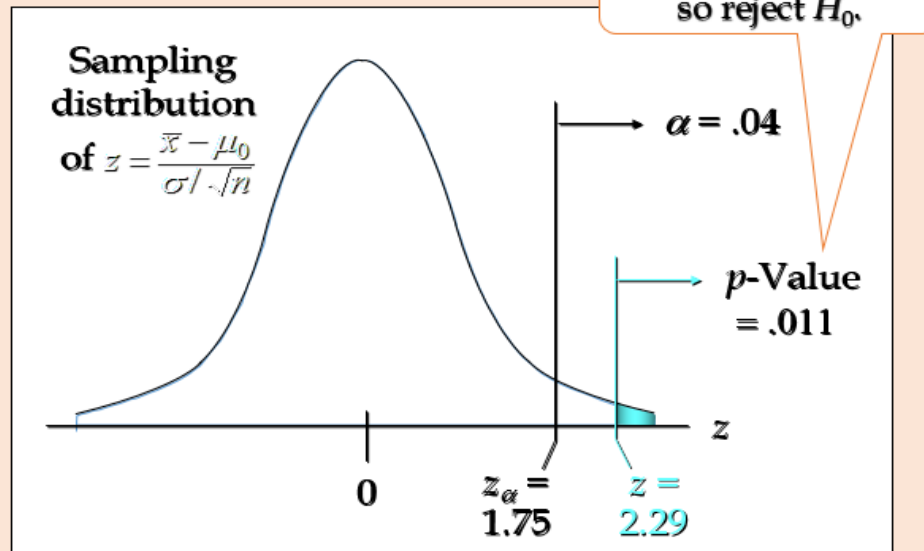
c. Pendekatan P-Value untuk Uji Satu Arah

- 1) Probabilitas dari pengamatan suatu nilai sampel yang sama ekstrem atau lebih ekstrem dari nilai yang diamati dengan hipotesis yang benar
- 2) Membandingkan probabilitas nilai p (P-value) dengan tingkat signifikansi
- 3) Jika nilai p lebih kecil daripada tingkat signifikansi, H_0 ditolak. Jika nilai p-value lebih besar dari tingkat signifikansi, maka H_0 diterima



Upper-Tailed Test About a Population Mean: σ Known

p-Value Approach

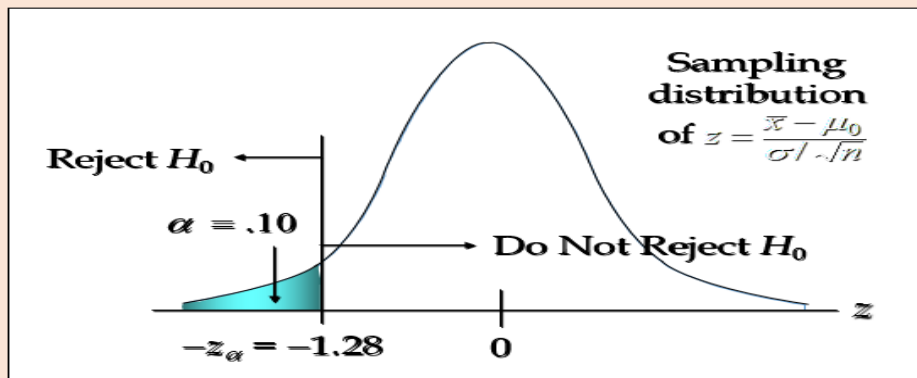


d. Pendekatan Critical Value Untuk Uji Hipotesis Satu Arah

- 1) Uji statistik z dengan menggunakan distribusi probabilitas normal standar
- 2) Kita dapat menggunakan tabel distribusi probabilitas normal standar untuk menemukan nilai-z dengan luas di bawah (atas) nilai distribusi.
- 3) Nilai statistik uji yang ditetapkan batas daerah penolakan disebut nilai kritis pengujian.
- 4) Tolah H_0 jika :
 - Lower Tail : $Z \leq -Z_\alpha$
 - Upper Tail : $Z \geq Z_\alpha$

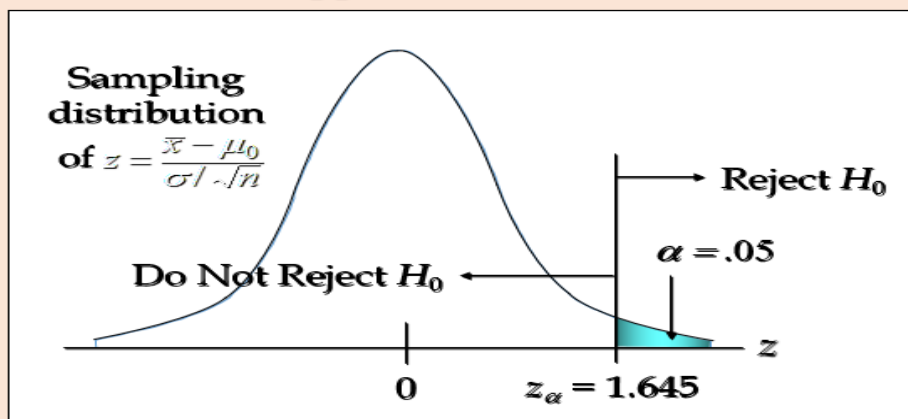
Lower-Tailed Test About a Population Mean: σ Known

Critical Value Approach



Upper-Tailed Test About a Population Mean: σ Known

Critical Value Approach



e. Langkah-Langkah Pengujian Hipotesis

Dengan pendekatan p-value :

- 1) Tentukanlah Hipotesis Nul (H_0) dan Hipotesis Alternatif (H_a)
- 2) Tentukanlah level of signifikan (tingkat signifikansi) α
- 3) Kumpulkan data sampel dan lakukan perhitungan statistik
- 4) Gunakan nilai statistik uji untuk menghitung p-value
- 5) Tolak H_0 , jika p-value $\leq \alpha$

Dengan Critical Value :

- 1) Tentukanlah Hipotesis Nul (H_0) dan Hipotesis Alternatif (H_a)
- 2) Tentukanlah level of signifikan (tingkat signifikansi) α

- 3) Kumpulkan data sampel dan lakukan perhitungan statistik
- 4) Gunakan tingkat significance α untuk menentukan nilai kritis dan aturan penolakan
- 5) Menggunakan nilai statistik uji dan aturan penolakan untuk menentukan apakah akan menolak H_0

f. Pendekatan P-Value Untuk Uji Dua Arah

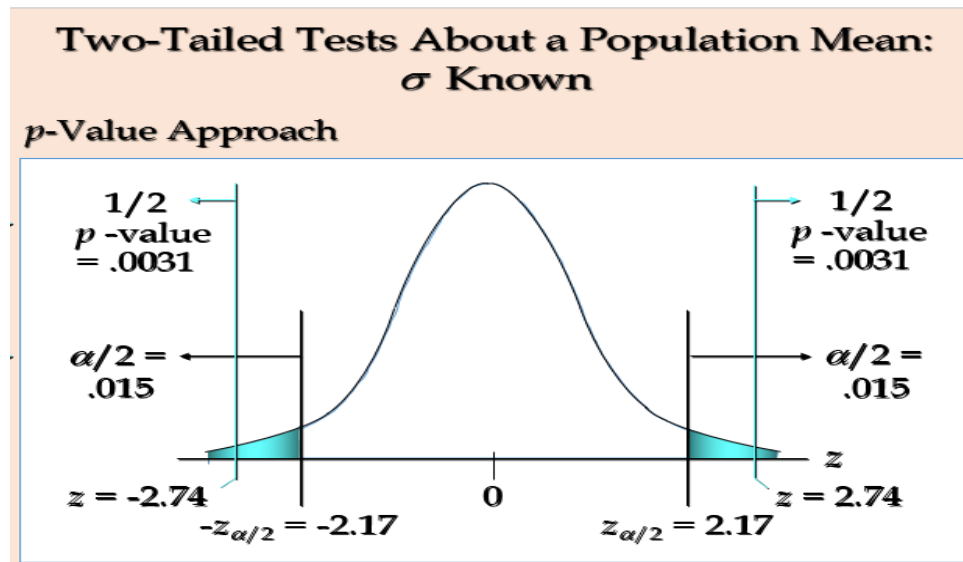
Langkah-langkah perhitungan P-Value :

- 1) Hitung nilai Z statistik uji
- 2) Jika z berada di bagian kanan ($z > 0$), menemukan area di bawah kurva normal standar kanan z. Jika z adalah di bagian kiri ($z < 0$), menemukan area di bawah kurva normal standar di sebelah kiri z.
- 3) Jika melakukan uji dua arah maka nilai P-Value dikalikan dua
- 4) Tolak H_0 : jika $p\text{-value} \leq \alpha$

g. Pendekatan Critical-Value Untuk Uji Dua Arah

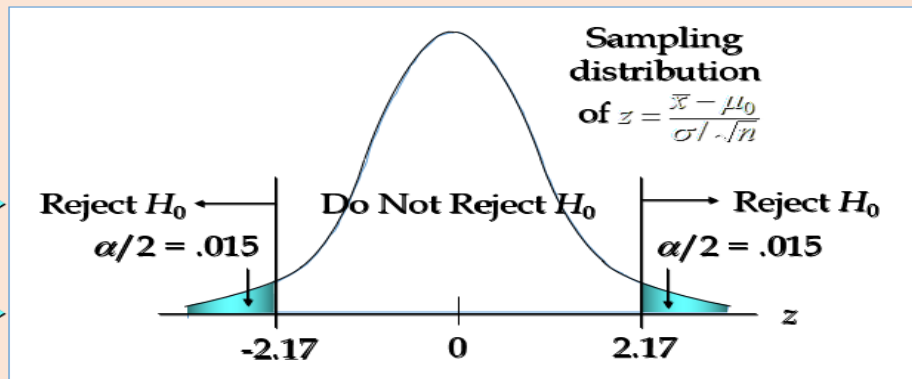
Langkah-langkah perhitungan P-Value :

- 1) Hitung nilai Z statistik uji
- 2) Dengan menggunakan kurva normal standar, dapatkan $Z_{\alpha/2}$ Jika z berada di bagian kanan ($z > 0$), menemukan area di bawah kurva normal standar kanan z. Jika z adalah di bagian kiri ($z < 0$), menemukan area di bawah kurva normal standar di sebelah kiri z.
- 3) Tolak H_0 : jika $Z \leq -Z_{\alpha/2}$ or $Z \geq Z_{\alpha/2}$



Two-Tailed Tests About a Population Mean: σ Known

Critical Value Approach



Uji Proporsi Populasi

| | | |
|-------------------------------------|-------------------------------------|-------------------------------------|
| $H_0: p \geq p_0$ $H_a: p < p_0$ | $H_0: p \leq p_0$ $H_a: p > p_0$ | $H_0: p = p_0$ $H_a: p \neq p_0$ |
| One-tailed (lower tail) | One-tailed (upper tail) | Two-tailed |

Tests About a Population Proportion

Test Statistic

$$z = \frac{\bar{p} - p_0}{\sigma_p}$$

where:

$$\sigma_p = \sqrt{\frac{p_0(1-p_0)}{n}}$$

assuming $np \geq 5$ and $n(1-p) \geq 5$

Tests About a Population Proportion

Rejection Rule: p -Value Approach

Reject H_0 if p -value $\leq \alpha$

Rejection Rule: Critical Value Approach

- ▶ $H_0: p \leq p_0$ Reject H_0 if $z \geq z_\alpha$
- ▶ $H_0: p \geq p_0$ Reject H_0 if $z \leq -z_\alpha$
- ▶ $H_0: p = p_0$ Reject H_0 if $z \leq -z_{\alpha/2}$ or $z \geq z_{\alpha/2}$

III. Uji Hipotesis dan Pengambilan Keputusan

- 1) In many decision-making situations the decision maker may want, and in some cases may be forced, to take action with both the conclusion do not reject H_0 and the conclusion reject H_0 .
- 2) In such situations, it is recommended that the hypothesis-testing procedure be extended to include consideration of making a Type II error.

a. Calculating the Probability of a Type II Error in Hypothesis Tests About a Population Mean

- 1) Formulasikan Hipotesis Nul dan Hipotesis Alternatif
- 2) Dengan menggunakan Critical Value pada tingkat signifikansi α , dapat digunakan untuk melakukan menolak hipotesis berdasarkan hasil uji Critical Value
- 3) Gunakan hasil uji statistik rata-rata untuk dibandingkan dengan nilai Critical Value, sebagai dasar pengambilan keputusan
- 4) Gunakan langkah ketiga untuk menentukan apakah menerima atau menolak H_0
- 5) Using the sampling distribution of \bar{x} for a value of μ satisfying the alternative hypothesis, and the acceptance region from step 4, compute the probability that the sample mean will be in the acceptance region. (This is the probability of making a Type II error at the chosen level of m .)

b. Determining the Sample Size for a Hypothesis Test About a Population Mean

Determining the Sample Size for a Hypothesis Test About a Population Mean

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_0 - \mu_a)^2}$$

where

z_α = z value providing an area of α in the tail

z_β = z value providing an area of β in the tail

σ = population standard deviation

μ_0 = value of the population mean in H_0

μ_a = value of the population mean used for the Type II error

Note: In a two-tailed hypothesis test, use $z_{\alpha/2}$ not z_α

Soal Latihan

1. Suatu pabrik susu merek Good Milk melakukan pengecekan terhadap produk mereka, apakah rata-rata berat bersih satu kaleng susu bubuk yang di produksi dan di pasaran masih tetap 400 gram atau sudah lebih kecil dari itu. Dari data sebelumnya di ketahui bahwa simpangan baku bersih per kaleng sama dengan 125 gram. Dari sample 50 kaleng yang di teliti, di peroleh rata-rata berat bersih 375 gram. Dapatkah di terima bahwa berat bersih rata-rata yang di pasaran tetap 400 gram? Ujilah dengan taraf nyata 5 % !

2. Seorang peneliti ingin mengetahui apakah catchability gillnet rata-rata masih tetap 30 ekor ikan atau lebih kecil dari itu. Data-data sebelumnya diketahui bahwa standard deviasi catchability 25 ekor. Sampel yang diambil 100 trip untuk diteliti dan diperoleh rata-rata tangkap 27 ekor. Apakah nilai tersebut masih dapat diterima sehingga catchability gillnet 30 ekor? Ujilah dengan taraf nyata 5%.

3. Populasi ikan lemuru hasil tangkapan purse seine panjang rata-rata 80 cm dengan standard deviasi 7 cm. Setelah 3 tahun beroperasi, konsumen meragukan panjang ikan tersebut. Guna meyakinkan keabsahan hipotesis itu, seorang peneliti mengambil sampel acak 100 ekor ikan lemuru dan diperoleh hasil perhitungan panjang rata-rata ikan adalah 83 cm dan standar deviasinya tetap.
Apakah ada alasan untuk meragukan bahwa rata-rata panjang ikan lemuru yang dihasilkan alat tangkap purse seine sama dengan 80 cm pada taraf signifikan 5% ?

Lembar Jawaban

| |
|--|
| |
| |
| |
| |
| |
| |
| |

Kriteria Penilaian:

- 4 : Apabila Skor latihan pertemuan 9, Uji Hipotesis 80 % > 80
- 3 : Apabila Skor latihan pertemuan 9, Uji Hipotesis 70 % : 70-79
- 2: Apabila Skor latihan pertemuan 9, Uji Hipotesis 50 % : 50-69
- 1: Apabila Skor latihan pertemuan 9, Uji Hipotesis 30 % : 31-59
- 0: Apabila Skor latihan pertemuan 9, Uji Hipotesis <30

| Diperiksa oleh Asisten Dosen | Di <i>review</i> Oleh Dosen |
|--|---|
| Soal 1: nilai 80 Soal 2: nilai 80 Soal 3: nilai 80 Soal 4: nilai 80 (contoh) | Tingkat ketercapaian pembelajaran tergolong tuntas. Berdasarkan skor latihan rata-rata 80. Skor 4 |

Statistik Inference Tentang Rata-Rata dan Proporsi Dengan 2 Populasi

Capaian Pembelajaran Perkuliahan :

1. Mampu menjelaskan dan mengerjakan statistik inference tentang rata-rata dan proporsi dengan 2 populasi
2. Mampu menjelaskan dan mengerjakan estimasi perbedaan rata-rata dua populasi sampel independen
3. Mampu melakukan uji hipotesis beda rata-rata antara dua populasi
4. Mampu menjelaskan dan melakukan perhitungan inference tentang variance populasi

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Statistik inference tentang rata-rata dan proporsi dengan 2 populasi
2. Estimasi perbedaan rata-rata dua populasi sampel independen
3. Uji hipotesis beda rata-rata antara 2 populasi
4. Inference tentang variance populasi

I. Statistik inference tentang rata-rata dan proporsi dengan 2 populasi

- a. Inference tentang perbedaan rata-rata dua populasi : σ_1 dan σ_2 diketahui
 - 1) Estimasi perbedaan antara rata-rata dua populasi
 - μ_1 adalah rata-rata populasi 1 dan μ_2 adalah rata-rata populasi 2
 - perbedaan dua populasi adalah $\mu_1 - \mu_2$
 - estimasi $\mu_1 - \mu_2$ dari sampel acak sederhana dengan ukuran sampel n_1 dari populasi 1, dan n_2 dari populasi 2
 - \bar{X}_1 adalah rata-rata sampel 1 dan \bar{X}_2 adalah rata-rata sampel 2
 - Poin estimasi adalah perbedaan rata-rata populasi 1 dan 2 $\bar{X}_1 - \bar{X}_2$

Sampling Distribution of $\bar{x}_1 - \bar{x}_2$

Expected Value

$$E(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$$

Standard Deviation (Standard Error)

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

where: σ_1 = standard deviation of population 1
 σ_2 = standard deviation of population 2
 n_1 = sample size from population 1
 n_2 = sample size from population 2

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Known

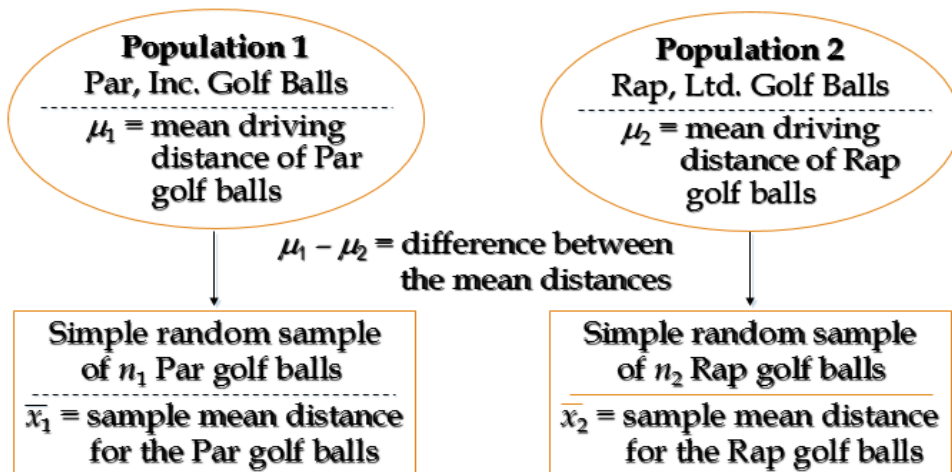
Interval Estimate

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

where:

$1 - \alpha$ is the confidence coefficient

Estimating the Difference Between Two Population Means



$\bar{x}_1 - \bar{x}_2$ = Point Estimate of $\mu_1 - \mu_2$

Hypothesis Tests About $\mu_1 - \mu_2$: σ_1 and σ_2 Known

Hypotheses

$$H_0: \mu_1 - \mu_2 \geq D_0$$

$$H_a: \mu_1 - \mu_2 < D_0$$

Left-tailed

$$H_0: \mu_1 - \mu_2 \leq D_0$$

$$H_a: \mu_1 - \mu_2 > D_0$$

Right-tailed

$$H_0: \mu_1 - \mu_2 = D_0$$

$$H_a: \mu_1 - \mu_2 \neq D_0$$

Two-tailed

Test Statistic

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Langkah-langkah Pengujian Hipotesis

- 1) Tentukan hipotesis nul dan hipotesis alternatif
- 2) Tentukan taraf signifikansi α
- 3) Hitung nilai uji statistik
- 4) Hitung nilai p-value
- 5) Bandingkan antara nilai p-value dengan nilai uji statistik, tentukan apakah menerima atau menolak H_0
- 6) Jika menggunakan critical value, bandingkan nilai critical value dengan nilai tabel normal

II. Inferences About the Difference Between Two Population Means: σ_1 and σ_2

Unknown

a. Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown

- 1) Jika σ_1 dan σ_2 tidak diketahui maka dipergunakan standar deviasi sampel s_1 dan s_2
- 2) Gunakan tabel t /2 sebagai dasar pengujian

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown

Interval Estimate

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Where the degrees of freedom for $t_{\alpha/2}$ are:

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

Point Estimate of $\mu_1 - \mu_2 = \bar{X}_1 - \bar{X}_2$

Hypothesis Tests About $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown

Hypotheses

| | | |
|-------------------------------|-------------------------------|-------------------------------|
| $H_0: \mu_1 - \mu_2 \geq D_0$ | $H_0: \mu_1 - \mu_2 \leq D_0$ | $H_0: \mu_1 - \mu_2 = D_0$ |
| $H_a: \mu_1 - \mu_2 < D_0$ | $H_a: \mu_1 - \mu_2 > D_0$ | $H_a: \mu_1 - \mu_2 \neq D_0$ |

Left-tailed

Right-tailed

Two-tailed

Test Statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Langkah-langkah Pengujian Hipotesis Untuk σ_1 dan σ_2 Tidak Diketahui

- 1) Tentukan hipotesis nul dan hipotesis alternatif
- 2) Tentukan level of signifikan α
- 3) Hitung nilai uji statistik
- 4) Dengan p-value tentukan apakah menerima atau menolak H_0

- 5) Dengan critical value bandingkan antara nilai uji statistik dengan nilai t tabel
- 6) Berikan kesimpulan

III. Inferences About the Difference Between Two Population Proportions

Sampling Distribution of $\bar{p}_1 - \bar{p}_2$

Expected Value

$$E(\bar{p}_1 - \bar{p}_2) = p_1 - p_2$$

Standard Deviation (Standard Error)

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

where: n_1 = size of sample taken from population 1
 n_2 = size of sample taken from population 2

Sampling Distribution of $\bar{p}_1 - \bar{p}_2$

If the sample sizes are large, the sampling distribution of $\bar{p}_1 - \bar{p}_2$ can be approximated by a normal probability distribution.

The sample sizes are sufficiently large if all of these conditions are met:

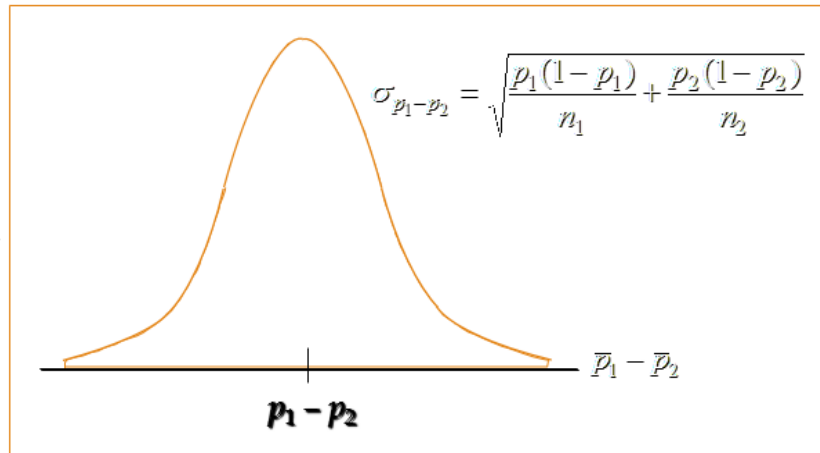
$$n_1 p_1 \geq 5$$

$$n_1(1 - p_1) \geq 5$$

$$n_2 p_2 \geq 5$$

$$n_2(1 - p_2) \geq 5$$

Sampling Distribution of $\bar{p}_1 - \bar{p}_2$



Interval Estimation of $p_1 - p_2$

Interval Estimate

$$\bar{p}_1 - \bar{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$$

Hypothesis Tests about $p_1 - p_2$

$$H_0: p_1 - p_2 \geq 0$$

$$H_a: p_1 - p_2 < 0$$

Left-tailed

$$H_0: p_1 - p_2 \leq 0$$

$$H_a: p_1 - p_2 > 0$$

Right-tailed

$$H_0: p_1 - p_2 = 0$$

$$H_a: p_1 - p_2 \neq 0$$

Two-tailed

Hypothesis Tests about $p_1 - p_2$

Standard Error of $\bar{p}_1 - \bar{p}_2$ when $p_1 = p_2 = p$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{p(1-p) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Pooled Estimator of p when $p_1 = p_2 = p$

$$\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2}$$

Hypothesis Tests about $p_1 - p_2$

Test Statistic

$$z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sqrt{\bar{p}(1-\bar{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Langkah-langkah Pengujian Hipotesis Beda Proporsi

- 1) Tentukan hipotesis nul dan hipotesis alternatif
- 2) Tentukan level of signifikan α
- 3) Hitung nilai uji statistik
- 4) Dengan p-value tentukan apakah menerima atau menolak H_0
- 5) Dengan critical value bandingkan antara nilai uji statistik dengan nilai t tabel
- 6) Berikan kesimpulan

Soal Latihan

1. Sebuah perusahaan mengadakan pelatihan teknik pemasaran. Sampel sebanyak 12 orang dengan metode biasa dan 10 orang dengan terprogram. Pada akhir pelatihan di berikan evaluasi dengan materi yang sama. Kelas pertama mencapai nilai rata-rata 75 dengan simpangan baku 4,5. Ujilah hipotesis kedua metode pelatihan, dengan alternative keduanya tidak sama! Gunakan taraf nyata 10%! Asumsikan kedua populasi menghampiri distribusi normal dengan varians yang sama!

Chapter 12

Goodness of Fit Tes dan Tes of Independence

Capaian Pembelajaran

1. Mahasiswa dapat melakukan uji kelayakan dan independensi model dengan menggunakan beberapa pendekatan

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Uji Goodness of Fit Tes : Populasi Multinomial
2. Uji Independensi
3. Uji Goodness of Fit Tes : Distribusi Normal

I. Uji Goodness of Fit Tes : Populasi Multinomial

Langkah-langkah Pengujian Goodness of Fit

- 1) Tentukan hipotesis non dan hipotesis alternatif
- 2) Seleksi sampel random catat frekuensi untuk setiap kategori
- 3) Hitung expected frekuensi e_i untuk setiap kategori
- 4) Hitung nilai statistik uji dengan rumus

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - e_i)^2}{e_i}$$

where:

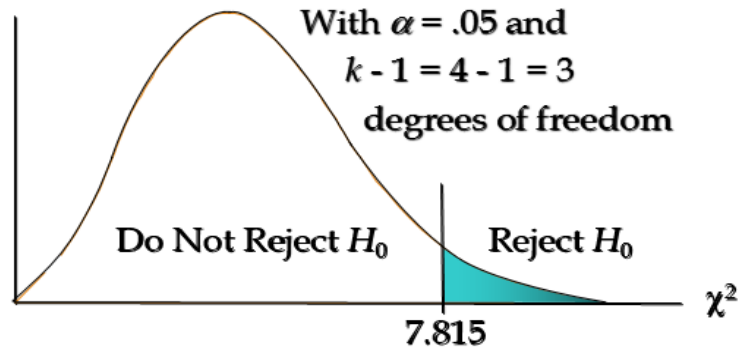
f_i = observed frequency for category i

e_i = expected frequency for category i

k = number of categories

Note: The test statistic has a chi-square distribution with $k - 1$ df provided that the expected frequencies are 5 or more for all categories.

- 5) Tolak H_0 jika : $p\text{-value} \leq \alpha$
Critical value $\chi \geq \chi^2$



II. Test of Independence: Contingency Tables

Langkah – langkah Tes Independence

- 1) Tentukan Hipotesis Nul dan Hipotesis Alternatif
- 2) Hitung frekuensi observasi f_{ij} untuk setiap sel
- 3) Hitung frekuensi expected e_{ij} untuk setiap sel, dengan rumus

$$e_{ij} = \frac{(\text{Row } i \text{ Total})(\text{Column } j \text{ Total})}{\text{Sample Size}}$$

- 4) Hitung nilai statistik uji dengan rumus :

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$$

- 5) Tolak H_0 jika :

$$\text{Reject } H_0 \text{ if } p\text{-value} \leq \alpha \text{ or } \chi^2 \geq \chi^2_{\alpha}$$

where α is the significance level and, with n rows and m columns, there are $(n - 1)(m - 1)$ degrees of freedom.

III. Goodness Of Fit : Distribusi Normal

Langkah-langkah Uji Goodness of Fit

- 1) Tentukan Hipotesis Nul dan Hipotesis Alternatif
- 2) A. hitung rata-rata dan standar deviasi
 - b. Hitung frekuensi expected paling sedikit 5 untuk setiap interval
 - c. Untuk setiap interval catat frekuensi observasi
- 3) Hitung expected frekuensi e_{ij} untuk setiap interval

4) Hitung nilai statistik uji dengan rumus :

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - e_i)^2}{e_i}$$

5) Tolak H_0 : jika $\chi^2 \geq \chi^2_{\alpha}$

Soal Latihan

1. Dalam pemilihan presiden lalu, sebelum para kandidat mulai kampanye utama mereka, persentase pemilih terdaftar yang disukai berbagai kandidat adalah sebagai berikut.

| | Percentages |
|--------------|-------------|
| Republicans | 34% |
| Democrats | 43% |
| Independents | 23% |

Dengan menggunakan sampel acak dari 400 pemilih menunjukkan bahwa 172 disukai kandidat Partai Republik; 164 yang mendukung kandidat Demokrat; dan 64 disukai calon Independen. Lembaga riset tertarik dalam menentukan apakah proporsi pemilih yang disukai berbagai kandidat telah berubah.

- Hitunglah uji statistiknya
 - Dengan menggunakan pendekatan p-value, apakah proporsi pemilih telah berubah
 - Dengan menggunakan critical value ujilah hipotesisnya
2. Selama beberapa minggu pertama musim televisi baru, proporsi penonton berita malam dicatat sebagai ABC 31%, CBS- 34%, dan 35% NBC-. Sebuah sampel dari 600 rumah menghasilkan data penonton berikut melihat.

| | Number of Homes |
|-----|-----------------|
| ABC | 150 |
| CBS | 200 |
| NBC | 250 |

Kita ingin menentukan apakah atau tidak telah terjadi perubahan signifikan dalam jumlah pemirsa dari tiga jaringan.

- Tentukanlah hipotesis nol dan alternatif .
 - Hitunglah frekuensi yang diharapkan.
 - Hitung statistik uji.
 - Dengan menggunakan tingkat kepercayaan 95%, ujilah hipotesisnya. Apa yang dapat disimpulkan.
3. Hasil penelitian terbaru mengenai merokok dan tiga jenis penyakit yang ditampilkan dalam tabel berikut.

Chapter 13

Analisis Variance

Capaian Pembelajaran

1. Mampu menjelaskan dan menggunakan Tabel ANOVA dalam penelitian untuk variabel dengan sampel acak dan populasi

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. ANOVA : Uji kesamaan k rata-rata Populasi
2. Multiple Comparison Prosedur

Analisis Variance (ANOVA)

- 1) Studi statistik dapat diklasifikasikan sebagai baik eksperimental atau observasional.
- 2) Dalam sebuah studi eksperimental, satu atau lebih faktor dikendalikan sehingga data dapat diperoleh tentang bagaimana faktor-faktor mempengaruhi variabel tersebut.
- 3) Dalam sebuah studi observasional, tidak ada usaha untuk mengendalikan faktor.
- 4) Hubungan sebab-akibat yang lebih mudah untuk dibangun dalam studi eksperimental daripada di studi observasional
- 5) Analisis varians (ANOVA) dapat digunakan untuk menganalisis data yang diperoleh dari studi eksperimental atau observasional.

I. ANOVA

a. Between-Treatments Estimate of Population Variance

- The estimate of σ^2 based on the variation of the sample means is called the mean square due to treatments and is denoted by MSTR.

$$MSTR = \frac{\sum_{j=1}^k n_j (\bar{x}_j - \bar{\bar{x}})^2}{k-1}$$

Denominator is the degrees of freedom associated with SSTR

Numerator is called the sum of squares due to treatments (SSTR)

b. Within-Treatments Estimate of Population Variance

- The estimate of σ^2 based on the variation of the sample observations within each sample is called the mean square error and is denoted by MSE.

$$MSE = \frac{\sum_{j=1}^k (n_j - 1) s_j^2}{n_T - k}$$

Denominator is the degrees of freedom associated with SSE

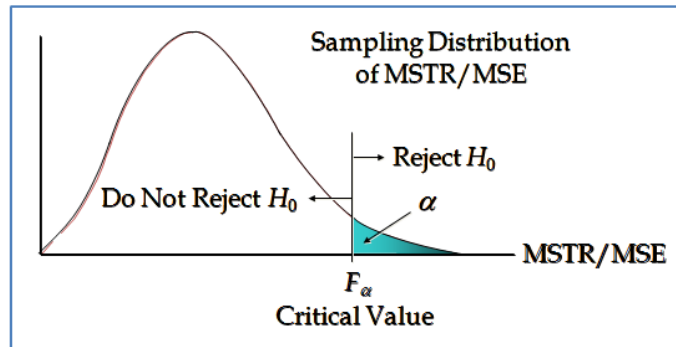
Numerator is called the sum of squares due to error (SSE)

c. Comparing the Variance Estimates: The F Test

- 1) Jika hipotesis nol adalah benar dan asumsi ANOVA berlaku, distribusi sampling dari MSTR / MSE adalah distribusi F dengan df MSTR sama dengan k - 1 dan MSE d.f. sama dengan nT - k.
- 2) Jika populasi k tidak sama, nilai MSTR / MSE akan meningkat karena MSTR overestimates α^2 .
- 3) Oleh karena itu, kita akan menolak H_0 jika nilai yang dihasilkan dari MSTR / MSE tampaknya terlalu besar dari distribusi F.

Comparing the Variance Estimates: The F Test

- Sampling Distribution of MSTR/MSE



d. ANOVA Table

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | F | p-Value |
|---------------------|----------------|--------------------|-----------------------------|--------------------|---------|
| Treatments | SSTR | k - 1 | $MSTR = \frac{SSTR}{k - 1}$ | $\frac{MSTR}{MSE}$ | |
| Error | SSE | $n_T - k$ | $MSE = \frac{SSE}{n_T - k}$ | | |
| Total | SST | $n_T - 1$ | | | |

SST is partitioned into SSTR and SSE.

SST's degrees of freedom (d.f.) are partitioned into SSTR's d.f. and SSE's d.f.

- 1) SST dibagi dengan derajat kebebasan $n_T - 1$ adalah varian sampel keseluruhan yang akan diperoleh jika kita diperlakukan seluruh set pengamatan sebagai satu set data.
- 2) Rumus untuk menghitung total jumlah kuadrat, SST, adalah:

$$SST = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 = SSTR + SSE$$

- 3) ANOVA dapat dilihat sebagai proses partisi total jumlah kotak dan derajat kebebasan dalam sumber yang berhubungan: perawatan dan kesalahan.
- 4) Membagi jumlah kuadrat dengan derajat kebebasan yang tepat memberikan perkiraan varians dan nilai F yang digunakan untuk menguji hipotesis populasi yang sama

e. Test for the Equality of k Population Means

- 1) Hipotesis

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

$$H_a: \text{Not all population means are equal}$$

- 2) Statistik Uji

$$F = MSTR/MSE$$

- 3) Tolak H_0 :

p -value Approach: $\text{Reject } H_0 \text{ if } p\text{-value} \leq \alpha$

Critical Value Approach: $\text{Reject } H_0 \text{ if } F \geq F_\alpha$

where the value of F_α is based on an F distribution with $k - 1$ numerator d.f. and $n_T - k$ denominator d.f.

f. Pengujian untuk Kesetaraan k Penduduk Berarti: Sebuah Studi observasional Pendekatan P-Value dan Critical Value

- 1) Tentukan hipotesis
- 2) Tentukan tingkat signifikansi α
- 3) Hitung nilai uji statistik
 - MSTR
 - MSE

Buat Tabel ANOVA

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | F | p-Value |
|---------------------|----------------|--------------------|-------------|------|---------|
| Treatment | 490 | 2 | 245 | 9.55 | .0033 |
| Error | 308 | 12 | 25.667 | | |
| Total | 798 | 14 | | | |

- 4) Hitung P-Value
- 5) Tolak H_0 : jika $p - \text{value} \leq 0.05$
- 6) Dengan Critical value : Tolak H_0 jika $F \geq F_\alpha$

II. Multiple Comparison Procedures

- 1) Misalkan analisis varian telah memberikan bukti statistik untuk menolak hipotesis nol dari mean populasi sama
- 2) Paling signifikan perbedaan (LSD) prosedur Fisher dapat digunakan untuk menentukan di mana perbedaan terjadi.

Fisher's LSD Procedure

- 1) Hipotesis

$$H_0 : \mu_i - \mu_j$$

$$H_a : \mu_i \neq \mu_j$$

- 2) Uji Statistik

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}$$

- 3) Tolak H_0 jika :

- P-value :

$$\text{Reject } H_0 \text{ if } p\text{-value} \leq \alpha$$

- Critical Value

$$\text{Reject } H_0 \text{ if } t < -t_{\alpha/2} \text{ or } t > t_{\alpha/2}$$

where the value of $t_{\alpha/2}$ is based on a t distribution with $n_T - k$ degrees of freedom.

Fisher's LSD Procedure Based on the Test Statistic $\bar{x}_i - \bar{x}_j$

Hypotheses

$$\begin{aligned} H_0 &: \mu_i - \mu_j \\ H_a &: \mu_i \neq \mu_j \end{aligned}$$

Test Statistic

$$\bar{x}_i - \bar{x}_j$$

Rejection Rule

$$\text{Reject } H_0 \text{ if } |\bar{x}_i - \bar{x}_j| > \text{LSD}$$

where

$$\text{LSD} = t_{\alpha/2} \sqrt{\text{MSE} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

Soal Latihan

- Informasi mengenai skor ACT sampel siswa di tiga jurusan yang berbeda diberikan di bawah ini.

| | Management | Major Finance | Accounting |
|-----------------|------------|------------------|------------|
| Sample size | 12 | 9 | 12 |
| Average | 24 | 25 | 26 |
| Sample variance | 18 | 7 | 10 |

- Hitunglah sampel mean keseluruhan.
 - Buat Tabel ANOVA .
 - Pada kepercayaan 95%, tentukan nilai kritis F.
 - Menggunakan pendekatan critical value, uji untuk menentukan apakah ada perbedaan yang signifikan dari tiga populasi.
 - Tentukan p-nilai dan apa kesimpulannya
- Gitar R. memiliki tiga toko yang terletak di tiga wilayah yang berbeda. Sampel acak dari penjualan tiga toko (dalam \$ 1000) adalah sebagai berikut.

| Store 1 | Store 2 | Store 3 |
|---------|---------|---------|
| 80 | 85 | 79 |
| 75 | 86 | 85 |
| 76 | 81 | 88 |
| 89 | 80 | |
| 80 | | |

- Hitunglah mean keseluruhan.
- Tentukanlah hipotesis nol dan alternatif yang akan diuji.

Chapter 14 & 15

Simple Linier Regression & Multiple Regression

Capaian Pembelajaran

1. Mampu menggunakan dan melakukan perhitungan dengan model regresi linier sederhana
2. Mampu menginterpretasikan hasil perhitungan model regresi linier berganda

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar Materi

1. Metode Least Square
2. Koefisien determinasi
3. Uji signifikansi
4. Penggunaan persamaan regresi untuk estimasi dan prediksi
5. Multiple coefficient of determinant
6. Uji signifikansi

I. Metode Least Square

- 1) Hubungan antara dua variabel atau lebih
- 2) Analisis regresi dapat digunakan untuk mengembangkan persamaan yang menunjukkan bagaimana variabel tersebut terkait satu sama lain
- 3) Variabel yang diprediksi disebut variabel dependen dan dilambangkan dengan y.
- 4) Variabel yang digunakan untuk memprediksi nilai variabel dependen disebut variabel independen dan dilambangkan dengan x.
- 5) Regresi linier sederhana melibatkan satu variabel independen dan satu variabel dependen
- 6) Persamaan yang menggambarkan bagaimana y berhubungan dengan x dan istilah error term disebut model regresi.
- 7) Model Regresi Linier Sederhana

$$y = \beta_0 + \beta_1 x + \varepsilon$$

where:

β_0 and β_1 are called parameters of the model,
 ε is a random variable called the error term.

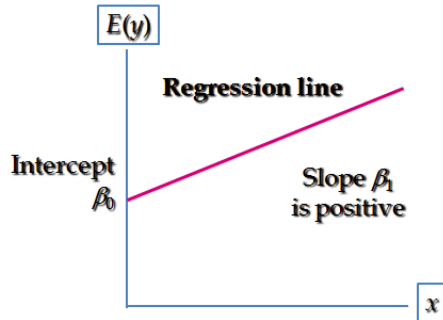
- 8) Persamaan Regresi Linier Sederhana :

$$E(y) = \beta_0 + \beta_1 x$$

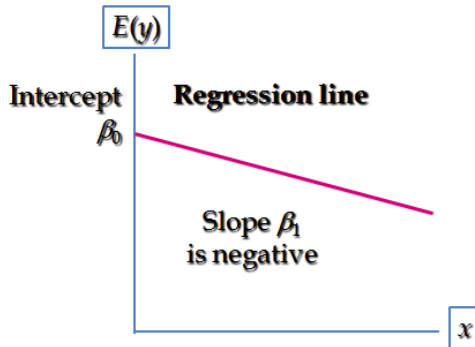
Graph of the regression equation is a straight line.
 β_0 is the y intercept of the regression line.
 β_1 is the slope of the regression line.
 $E(y)$ is the expected value of y for a given x value.

9) Hubungan Kedua Variabel

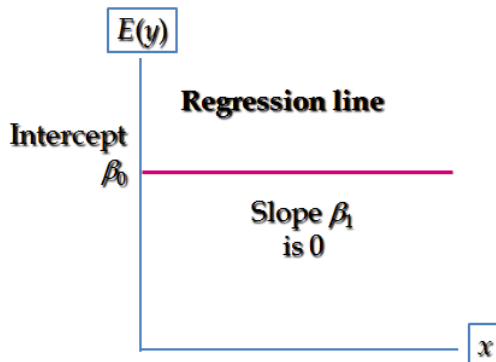
a. Hubungan Positif



b. Hubungan Negatif



c. Tidak ada hubungan

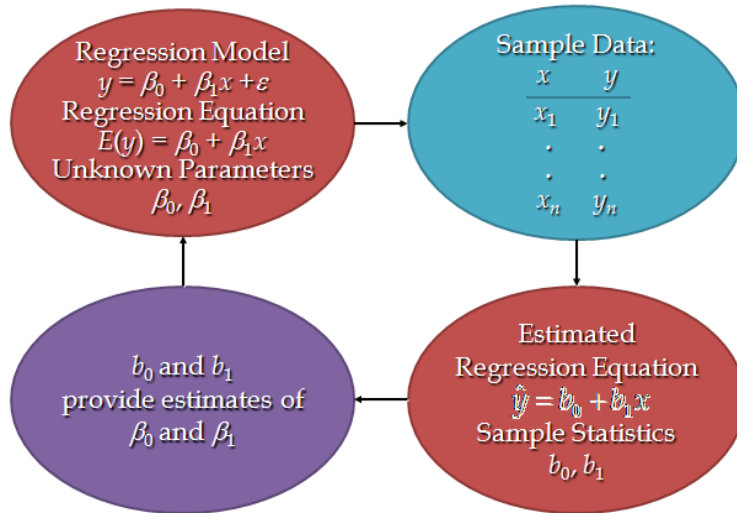


10) Estimasi Persamaan Linier Sederhana

$$\hat{y} = b_0 + b_1x$$

- The graph is called the estimated regression line.
- b_0 is the y intercept of the line.
- b_1 is the slope of the line.
- \hat{y} is the estimated value of y for a given x value.

Process Estimasi



11) Metode Least Square (Kuadrat Terkecil)

Least Squares Criterion

$$\min \sum (y_i - \hat{y}_i)^2$$

where:

y_i = observed value of the dependent variable for the i th observation

\hat{y}_i = estimated value of the dependent variable for the i th observation

Slope for the Estimated Regression Equation

$$b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

where:

x_i = value of independent variable for i th observation

y_i = value of dependent variable for i th observation

\bar{x} = mean value for independent variable

\bar{y} = mean value for dependent variable

y-Intercept for the Estimated Regression Equation

$$b_0 = \bar{y} - b_1 \bar{x}$$

II. Koefisien Determinasi

Relationship Among SST, SSR, SSE

$$SST = SSR + SSE$$

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2$$

where:

SST = total sum of squares

SSR = sum of squares due to regression

SSE = sum of squares due to error

The coefficient of determination is:

$$r^2 = SSR/SST$$

where:

SSR = sum of squares due to regression

SST = total sum of squares

Sample Correlation Coefficient

$$r_{xy} = (\text{sign of } b_1) \sqrt{\text{Coefficient t of Determination}}$$

$$r_{xy} = (\text{sign of } b_1) \sqrt{r^2}$$

where:

b_1 = the slope of the estimated regression equation $\hat{y} = b_0 + b_1 x$

Koefisien Korelasi

$$r_{xy} = (\text{sign of } b_1) \sqrt{r^2}$$

a. Asumsi Error Term

- 1) Kesalahan ϵ adalah variabel random dengan mean nol
- 2) Varians dari ϵ , dinotasikan dengan σ^2 , adalah sama untuk semua nilai variabel independen.
- 3) Nilai ϵ adalah independen
- 4) Error ϵ adalah terdistribusi normal

III. Uji Signifikasi

- 1) Untuk menguji hubungan regresi yang signifikan, kita harus melakukan pengujian hipotesis untuk menentukan apakah nilai b_1 adalah nol.
- 2) Dua tes yang sering dipergunakan



- 3) Kedua uji t dan uji F membutuhkan perkiraan \hat{s}^2 , varians dari ϵ dalam model regresi

An Estimate of σ^2

The mean square error (MSE) provides the estimate of σ^2 , and the notation s^2 is also used.

$$s^2 = \text{MSE} = \text{SSE} / (n - 2)$$

where:

$$\text{SSE} = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - b_0 - b_1 x_i)^2$$

$$s = \sqrt{\text{MSE}} = \sqrt{\frac{\text{SSE}}{n - 2}}$$

- 4) Uji Sigifikasi : Uji t
Hypotheses

$$\begin{aligned} H_0: \beta_1 &= 0 \\ H_a: \beta_1 &\neq 0 \end{aligned}$$

Test Statistic

$$t = \frac{b_1}{s_{b_1}}$$

where

$$s_{b_1} = \frac{s}{\sqrt{\sum (x_i - \bar{x})^2}}$$

Tolak H_0 , jika :

Reject H_0 if $p\text{-value} \leq \alpha$
 or $t \leq -t_{\alpha/2}$ or $t \geq t_{\alpha/2}$

where:

$t_{\alpha/2}$ is based on a t distribution
 with $n - 2$ degrees of freedom

5) Uji Signifikasi : Uji F

Hypotheses

$H_0: \beta_1 = 0$
 $H_a: \beta_1 \neq 0$

Test Statistic

$F = MSR/MSE$

Tolak H_0 , jika :

Reject H_0 if
 $p\text{-value} \leq \alpha$
 or $F \geq F_\alpha$

where:

F_α is based on an F distribution with
 1 degree of freedom in the numerator and
 $n - 2$ degrees of freedom in the denominator

IV. Penggunaan Persamaan Regresi Untuk Estimasi dan Prediksi

Confidence Interval Estimate of $E(y_p)$

$\hat{y}_p \pm t_{\alpha/2} s_{\hat{y}_p}$

Prediction Interval Estimate of y_p

$\hat{y}_p \pm t_{\alpha/2} s_{ind}$

where:

confidence coefficient is $1 - \alpha$ and
 $t_{\alpha/2}$ is based on a t distribution
 with $n - 2$ degrees of freedom

$$s_{\hat{y}_p} = s \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$$

V. Model Regresi Berganda

- 1) Persamaan yang menggambarkan bagaimana variabel dependen y berhubungan dengan variabel independen x_1, x_2, \dots, x_p dan error term
- 2) Model persamaan regresi :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon$$

where:

$\beta_0, \beta_1, \beta_2, \dots, \beta_p$ are the **parameters**, and ε is a **random variable called the error term**

- 3) Persamaan Regresi Berganda :

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

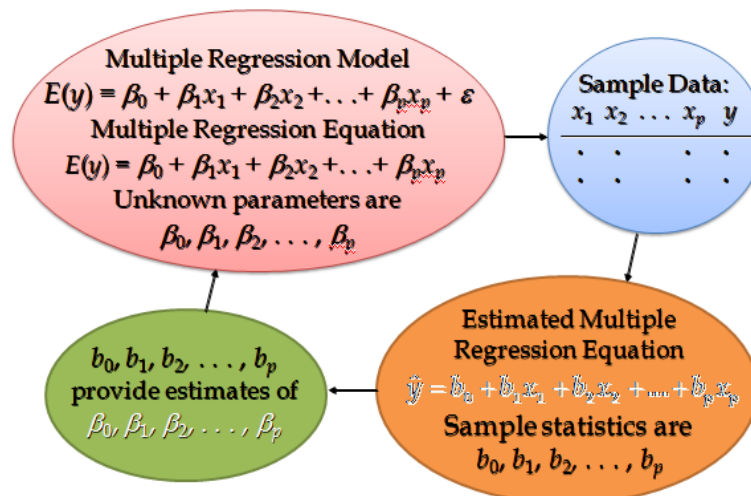
- 4) Estimasi Persamaan Regresi Berganda

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p$$

A simple random sample is used to compute sample statistics $b_0, b_1, b_2, \dots, b_p$ that are used as the point estimators of the parameters $\beta_0, \beta_1, \beta_2, \dots, \beta_p$.

- 5) Proses Estimasi

Proses Estimasi



6) Metode Kuadrat Terkecil (Least Square Methods)

$$\min \sum (y_i - \hat{y}_i)^2$$

7) Koefisien Determinasi Regresi berganda

$$\begin{aligned} \text{SST} &= \text{SSR} + \text{SSE} \\ \sum (y_i - \bar{y})^2 &= \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2 \end{aligned}$$

where:

- SST = total sum of squares
- SSR = sum of squares due to regression
- SSE = sum of squares due to error

8) Koefisien Determinasi Yang Telah Disesuaikan

$$R_a^2 = 1 - (1 - R^2) \frac{n - 1}{n - p - 1}$$

Asumsi Error Term

- 1) Kesalahan ϵ adalah variabel random dengan mean nol
 - 2) Varians dari ϵ , dinotasikan dengan σ^2 , adalah sama untuk semua nilai variabel independen.
 - 3) Nilai ϵ adalah independen
 - 4) Error ϵ adalah terdistribusi normal dari nilai $b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p$.
- 9) Uji Signifikasi
 Uji keseluruhan : Uji F
 Uji individual : Uji t

Soal Latihan

1. Asumsikan Anda telah mencatat harga berikut untuk buku dan jumlah halaman yang masing-masing buku berisi.

| Book | Pages (x) | Price (y) |
|------|-----------|-----------|
| A | 500 | \$7.00 |
| B | 700 | 7.50 |
| C | 750 | 9.00 |
| D | 590 | 6.50 |
| E | 540 | 7.50 |
| F | 650 | 7.00 |
| G | 480 | 4.50 |

- a. Gunakan metode kuadrat terkecil, tentukan persamaan regresinya

- b. Hitunglah koefisien determinasi dan jelaskan maknanya.
- c. Hitunglah koefisien korelasi antara harga dan jumlah halaman. Tes untuk melihat apakah x dan y terkait. Gunakan $\alpha = 0,10$.
2. Data berikut merupakan jumlah flash drive yang dijual per hari di sebuah toko komputer lokal dan harga mereka

| Price (x) | Units Sold (y) |
|-----------|----------------|
| \$34 | 3 |
| 36 | 4 |
| 32 | 6 |
| 35 | 5 |
| 30 | 9 |
| 38 | 2 |
| 40 | 1 |

- a. Gunakan metode kuadrat-terkecil, tentukan persamaan regresi dan jelaskan kemiringan garis yang didapatkan
- b. Hitunglah koefisien determinasi dan interpretasikan kekuatan hubungan antara x dan y.
- c. Hitunglah koefisien korelasi sampel antara harga dan jumlah flash drive dijual. Gunakan $\alpha = 0,01$ untuk menguji hubungan antara x dan y.
3. Analisis regresi berganda digunakan untuk mempelajari bagaimana pendapatan individu (Y dalam ribuan dolar) dipengaruhi oleh usia (X1 dalam tahun), tingkat pendidikan (X2 mulai dari 1 sampai 5), dan jenis kelamin orang tersebut (X3 di mana 0 = perempuan dan 1 = laki-laki). Berikut ini adalah hasil parsial dari program komputer yang digunakan pada sampel dari 20 individu.

| | Coefficient | Standard Error |
|----------------|-------------|----------------|
| X ₁ | 0.6251 | 0.094 |
| X ₂ | 0.9210 | 0.190 |
| X ₃ | -0.510 | 0.920 |

Analysis of Variance

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square | F |
|---------------------|--------------------|----------------|-------------|---|
| Regression | | 84 | | |
| Error | | 112 | | |

- a. Hitunglah koefisien determinasi.

| |
|--|
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |

| Lembar Evaluasi | |
|---|--|
| Kriteria Penilaian: 4 : Apabila Skor latihan pertemuan 13, Simple & Multiple Regression 80 % > 80 3 : Apabila Skor latihan pertemuan 13, Simple & Multiple Regression 70 %: 70-79 2: Apabila Skor latihan pertemuan 13, Simple & Multiple Regression 50 %: 50-69 1: Apabila Skor latihan pertemuan 13, Simple & Multiple Regression 30 % : 31-59 0: Apabila Skor latihan pertemuan 13, Simple & Multiple Regression <30 | |
| Diperiksa oleh Asisten Dosen | Di <i>review</i> Oleh Dosen |
| Soal 1: nilai 80 Soal 2: nilai 80 Soal 3: nilai 80 Soal 4: nilai 80 (contoh) | Tingkat ketercapaian pembelajaran tergolong tuntas. Berdasarkan skor latihan rata-rata 80. Skor 4 |

Chapter 19

Statistik Non Parametrik

Capaian Pembelajaran

1. Mampu menggunakan dan melakukan perhitungan statistik non parametrik dalam penelitian

Referensi

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, 2012., : Statistics for Business and Economics, 11th Edition, South-Western International Edition (AS)

Ikhtisar materi

1. Sign test with small sample case dan large sample case
2. Wilcoxon signed rank test
3. Man-Whitney – Wilcoxon ttt
4. Kruskal walis test
5. Rank correlation

Statistik Non Parametrik :

- 1) Sebagian besar metode statistik disebut sebagai parametrik memerlukan penggunaan data interval- atau rasio-skala.
- 2) Metode nonparametrik sering satu-satunya cara untuk menganalisis data nominal atau ordinal dan menarik kesimpulan statistik.
- 3) Metode nonparametrik tidak memerlukan asumsi tentang distribusi probabilitas populasi
- 4) Metode nonparametrik sering disebut metode distribusi bebas.

I. Sign test with small sample case dan large sample case

- 1) Sebuah aplikasi umum dari uji tanda melibatkan menggunakan sampel n pelanggan potensial untuk mengidentifikasi preferensi untuk salah satu dari dua merek suatu produk.
- 2) Tujuannya adalah untuk menentukan apakah ada perbedaan preferensi antara dua item yang dibandingkan.
- 3) Untuk merekam data preferensi, kita menggunakan tanda plus jika individu lebih suka satu brand dan tanda minus jika individu lebih memilih merek lain

4) Karena data dicatat sebagai plus dan minus tanda-tanda, tes ini disebut uji tanda.

5) Uji Tanda untuk : Kasus sampel Kecil

- Kasus sampel kecil untuk uji tanda harus digunakan setiap kali $n < 20$.

- Hipotesis :

$$H_0 : p = 0.50$$

$$H_a : p \neq 0.50$$

- Tolak H_0 , jika : P-Value \leq level of significant α
- Jika sampel lebih besar ($n > 20$) gunakan pendekatan normal
- Mean dan Standar Deviasi

$$\text{Mean : } \mu = 0.50n$$

$$\text{Standar Deviasi : } \sigma = \sqrt{0.25n}$$

- Statistik uji

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

- Tolak H_0 , jika : p-value \leq level of signifikan α

6) Untuk kasus sampel besar

- Menggunakan tanda plus setiap kali data dalam sampel berada di atas nilai hipotesis dari median
- Menggunakan tanda minus setiap kali data dalam sampel berada di bawah nilai hipotesis dari median
- Membuang data persis sama dengan median hipotesis

II. Wilcoxon Signed-Rank Test

1) Metode ini membutuhkan data interval

2) Asumsi bahwa populasi perbedaan antara pasang pengamatan terdistribusi secara normal.

3) Langkah-langkah perhitungan :

- a. Hitung perbedaan antara pasangan pengamatan
- b. Buang angka nol
- c. Peringkat nilai absolut perbedaan dari terendah ke tertinggi. Perbedaan terikat ditugaskan peringkat rata-rata posisi mereka.
- d. Berikan angka mutlak dari perbedaan
- e. Jumlahkan tandanya

4) Rumus Statistik Uji :

$$\text{Mean} = \mu_T = 0$$

$$\text{Standar Deviasi : } \sigma_T = \frac{\sqrt{n(n+1)(2n+1)}}{6}$$

$$Z = \frac{T - \mu_T}{\sigma_T}$$

- 5) Uji hipotesis :
Tolak H_0 , jika : p-value \leq level of signifikan α

III. Man Whitney Wilcoxon Test

- 1) Metode ini adalah metode lain nonparametrik untuk menentukan apakah ada perbedaan antara dua populasi
- 2) Tidak seperti Wilcoxon test, metode ini tidak didasarkan pada kecocokan sampel
- 3) Uji ini tidak memerlukan data interval atau asumsi bahwa kedua populasi berdistribusi normal.
- 4) Satu-satunya persyaratan adalah bahwa skala pengukuran untuk data tersebut setidaknya ordinal.
- 5) Metode uji ini untuk menentukan apakah dua populasi identik, dengan persyaratan adalah bahwa skala pengukuran untuk data setidaknya ordinal.
- 6) Hipotesis :
 H_0 : dua populasi adalah identik
 H_a : dua populasi adalah tidak identik
- 7) Langkah-langkah pengujian :
 - a. Beri peringkat data gabungan dari yang terendah ke nilai tertinggi
 - b. Hitung T, untuk jumlah peringkat dari sampel pertama
 - c. Kemudian, membandingkan nilai yang diamati dari T untuk distribusi sampling dari T untuk populasi yang identik. Nilai standar uji statistik z akan memberikan dasar untuk memutuskan apakah akan menolak H_0 .

Mann-Whitney-Wilcoxon Test: Large-Sample Case

Sampling Distribution of T for Identical Populations

- Mean

$$\mu_T = 1/2 n_1(n_1 + n_2 + 1)$$

- Standard Deviation

$$\sigma_T = \sqrt{1/12 n_1 n_2 (n_1 + n_2 + 1)}$$

- Distribution Form

Approximately normal, provided
 $n_1 \geq 10$ and $n_2 \geq 10$

- d. Uji hipotesis :

$$Z = \frac{T - \mu_T}{\sigma_T}$$

e. Tolak H_0 , jika : p-value \leq level of signifikan α

IV. Kruskal Wallis Test

- 1) Uji Kruskal Wallis dipergunakan untuk kasus tiga atau lebih populasi.
- 2) Hipotesisnya
 H_0 : seluruh populasi adalah identik
 H_a : tidak seluruh populasi adalah identik
- 3) Uji Kruskal-Wallis dapat digunakan dengan data ordinal serta dengan data interval atau rasio.
- 4) Uji Kruskal Wallis tidak memerlukan asumsi data berdistribusi normal
- 5) Uji statistik

$$W = \left[\frac{12}{n_T(n_T + 1)} \sum_{i=1}^k \frac{R_i^2}{n_i} \right] - 3(n_T + 1)$$

where: k = number of populations

n_i = number of items in sample i

$n_T = \sum n_i$ = total number of items in all samples

R_i = sum of the ranks for sample i

- 6) Ketika populasi identik, distribusi sampling dari statistik uji W dapat didekati dengan distribusi chi-kuadrat dengan $k - 1$ derajat kebebasan.
- 7) Ketika populasi identik, pengambilan sampel distribusi dari statistik uji W dapat didekati dengan distribusi chi-kuadrat dengan derajat kebebasan $k - 1$
- 8) Tolak H_0 , jika : p-value \leq level of signifikan α

V. Rank Spearman Test

- 1) Koefisien korelasi Pearson, r , adalah ukuran dari asosiasi linear antara dua variabel data interval atau rasio
- 2) Koefisien rank Spearman rank, (r_s), adalah ukuran dari hubungan antara dua variabel jika hanya data ordinal yang tersedia.
- 3) Nilai-nilai r_s dapat berkisar dari -1.0 ke +1.0, dimana :
 - a. nilai mendekati 1.0 menunjukkan hubungan positif yang kuat, dan
 - b. nilai mendekait -1,0 menunjukkan hubungan negatif yang kuat.
- 4) Rumus korelasi Rank spearman

■ Spearman Rank-Correlation Coefficient, r_s

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

where: n = number of items being ranked

x_i = rank of item i with respect to one variable

y_i = rank of item i with respect to a second variable

$d_i = x_i - y_i$

5) Uji hipotesis

$H_0 : \rho_s = 0$ (No rank correlation exists)

$H_a : \rho_s \neq 0$ (Rank correlation exists)

Sampling Distribution of r_s when $\rho_s = 0$

- Mean

$$\mu_{r_s} = 0$$

- Standard Deviation

$$\sigma_{r_s} = \sqrt{\frac{1}{n-1}}$$

- Distribution Form

Approximately normal, provided $n \geq 10$

Soal Latihan

1. Dalam sampel 400 orang, 250 menunjukkan bahwa mereka lebih suka anggur domestik, sementara 140 mengatakan mereka lebih suka anggur Eropa, dan 10 menunjukkan tidak ada preferensi. Kami ingin menggunakan uji tanda untuk menentukan apakah ada bukti perbedaan yang signifikan dalam preferensi untuk dua jenis anggur.
 - a. Tentukan hipotesis yang akan diuji.
 - b. Hitunglah mean.
 - c. Menghitung standar deviasi.
 - d. Pada kepercayaan 95%, uji untuk menentukan apakah ada bukti perbedaan yang signifikan dalam preferensi untuk dua jenis anggur.
2. Catatan penjualan bulanan dari dua cabang dari sebuah department store (Cabang A dan B Cabang) selama setahun terakhir (12 bulan) yang dikumpulkan. Diputuskan untuk menggunakan uji Mann-Whitney-Wilcoxon untuk menentukan apakah telah terjadi perbedaan yang signifikan antara penjualan dua cabang.
 - a. Tentukan hipotesis untuk tes ini.
 - b. Hitunglah mean μ_T .

