

**Buku Ajar**  
**STATISTIKA BISNIS**  
TEORI DAN APLIKASI PROGRAM R  
(Paket R-Commander)

**Azwar Habibi**



# BUKU AJAR STATISTIKA BISNIS

© xiv+242; 16x24 cm

Juli 2021

Penulis : Azwar Habibi, S.Si., M.Si.

Editor : Fatati Nuryana, M.Si.

Khotibul Umam, M.E.I.

Layout &

Desain Cover : Duta Creative

## Duta Media Publishing

Jl. Masjid Nurul Falah Lekoh Barat Bangkes Kadur Pamekasan, Call/WA:  
082 333 061 120, E-mail: [redaksi.dutamedia@gmail.com](mailto:redaksi.dutamedia@gmail.com)

*All Rights Reserved.*

Dilarang memperbanyak sebagian atau seluruh isi buku ini dalam bentuk  
apa pun tanpa ijin tertulis dari penerbit

ISBN: 978-623-6705-82-7

IKAPI: 180/JTI/2017

**Undang-Undang Republik Indonesia  
Nomor 19 tahun 2002  
Tentang Hak Cipta**

**Lingkup Hak Cipta**

**Pasal 2**

1. Hak Cipta merupakan hak eksklusif bagi Pencipta atau Pemegang Hak Cipta untuk mengumumkan atau memperbanyak Ciptaannya, yang timbul secara otomatis setelah suatu ciptaan dilahirkan tanpa mengurangi pembatasan menurut peraturan perundang-undangan yang berlaku.

**Ketentuan Pidana**

1. Barangsiapa dengan sengaja dan tanpa hak melakukan perbuatan sebagaimana dimaksud dalam Pasal 2 ayat (1) atau Pasal 49 ayat (1) dan ayat (2) dipidana dengan pidana penjara masing-masing paling singkat 1 (satu) bulan dan/atau denda paling sedikit Rp 1.000.000,00 (satu juta rupiah), atau pidana penjara paling lama 7 (tujuh) tahun dan/atau denda paling banyak Rp. 5.000.000.000,00 (lima miliar rupiah).
2. Barangsiapa dengan sengaja menyiarkan, memamerkan, mengedarkan, atau menjual kepada umum suatu Ciptaan atau barang hasil pelanggaran Hak Cipta atau Hak Terkait sebagaimana dimaksud pada ayat (1) dipidana dengan pidana penjara paling lama 5 (lima) tahun dan/atau denda paling banyak Rp 500.000.000,00 (lima ratus juta rupiah).

# HALAMAN PENGESAHAN

## LEMBAR IDENTITAS DAN PENGESAHAN

1. Judul Buku Ajar : STATISTIK BISNIS (TEORI DAN APLIKASI PROGRAM R (PAKET R-Commander))
2. Mata Kuliah : STATISTIKA BISNIS
3. Nama Penulis : AZWAR HABIBI, S.Si. M.Si
4. Nomor Induk Pegawai : 198504242020121008
5. Fakultas : Fakultas Ekonomi dan Bisnis Islam

Pamekasan, Juli 2021

Reviewer

Penulis

**Fatati Nuryana, S.Si., M.Si**

NIP. 197910012005012004

**Azwar Habibi, S.Si., M.Si**

NIP. 198504242020121008

Mengetahui  
Dekan Fakultas Ekonomi  
Bisnis Islam

Mengesahkan  
Rektor IAIN Madura

**Dr. Zainal Abidin M.E.I.**

NIP. 198007012006041005

**Dr. H. Mohammad Kosim, M.Ag**

NIP.196901011994031008

## **Kata Sambutan**

*Assalamu'alaikum Wr. Wb.* Alhamdulillah, segala puji dan syukur kita panjatkan kehadirat Allah SWT atas segala rahmat, taufiq dan hidayah-Nya kepada kita semua sehingga penyusunan buku ajar ini dapat terselesaikan dengan baik. Shalawat dan salam senantiasa kita sanjungkan kepada junjungan kita Nabi Besar Muhammad SAW yang telah menuntun kita sebagai muslim ke jalan yang lurus dan diridhai Allah SWT. Menulis buku ajar ini sampai pada tahap cetak, merupakan bagian akademik Dosen Fakultas Ekonomi dan Bisnis Islam IAIN Madura yang harus diselesaikan untuk memperoleh Tenaga Edukatif. Untuk memperlancar dan menyeragamkan proses kegiatannya, diperlukan adanya buku ajar dan, bagi para dosen dilingkungan Fakultas Ekonomi dan Bisnis Islam. Buku Ajar ini terdiri dari tiga belas bab. Seluruh mahasiswa harus berpedoman pada buku ajar ini dalam melaksanakan proses ujian perkuliahan. Kami menyadari sepenuhnya bahwa buku ajar ini masih belum sempurna. Berbagai saran dan kritik konstruktif sangat kami harapkan untuk melakukan revisi dan perbaikan pada tahap-tahap selanjutnya. Buku ajar ini diharapkan dapat memberikan manfaat bagi semua pihak yang memerlukan, baik dalam rangka melaksanakan tugas akademik di Fakultas Ekonomi dan Bisnis Islam IAIN Madura maupun tugas-tugas lainnya. Akhirnya, kami mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang terlibat dalam penyusunan buku ajar ini. Kami juga mengucapkan terima kasih kepada penulis yang telah banyak meluangkan waktu, mencurahkan tenaga dan pikirannya, hingga tersusunnya buku ajar ini. *Wassalamu'alaikum Wr. Wb.*

Dekan Fakultas Ekonomi Bisnis Islam

**Dr. Zainal Abidin M.E.I.**

# Kata Pengantar

Alhamdulillah, segala puji dan syukur Reviewer panjatkan kehadirat Allah SWT karena buku ini telah selesai disusun. Buku ini disusun agar dapat membantu para mahasiswa dalam mempelajari konsep-konsep Statistika Bisnis beserta mempermudah mempelajari materi konsep dan Aplikasi menggunakan Software dalam Statistika Bisnis terutama bagi kaum awam yang belum mengenal Statistika Bisnis itu sendiri.

Reviewer pun menyadari jika didalam penyusunan buku ini mempunyai kekurangan, namun penulis meyakini sepenuhnya bahwa sekecil apapun buku ini tetap akan memberikan sebuah manfaat bagi pembaca.

Akhir kata untuk penyempurnaan buku ini, maka kritik dan saran dari pembaca sangatlah berguna untuk penulis kedepannya.

Reviewer

**Fatati Nuryana, S.Si., M.Si**

# Prakata

Puji syukur dipanjatkan ke hadirat Tuhan yang Maha Esa, atas rahmat-Nya Buku Ajar Statistika Bisnis ini dapat diselesaikan sesuai jadwal yang direncanakan. Tujuan utama penyusunan buku ajar ini adalah untuk membantu para mahasiswa mempelajari Statistika Bisnis secara lebih mudah dan terstruktur. Untuk memudahkan mahasiswa, dalam Buku ajar ini diberikan teori konsep dan contoh kasus serta penyelesaian menggunakan bantuan program R menggunakan Paket R-Commander, agar mahasiswa dapat mengetahui kejelasan langkah analisis datanya.

Buku ajar ini terdiri dari 13 Bab, yaitu :

Bab I. Konsep Dasar Statistika

Bab II. Program Statistika R

Bab III. Statistika Deskriptif

Bab IV. Pengujian Hipotesis

Bab V. Pengujian Data

Bab VI. Uji Beda Rata-Rata

Bab VII. Pengujian Hipotesis Tentang Proporsi

Bab VIII. Tabel Kontingensi

Bab IX. Pengujian Instrumen Penelitian

Bab X. Statistika Non Parametrik

Bab XI. Analisis Korelasi

Bab XII. Analisis Regresi

Bab XIII. Metode Data Envelopment Analysis (DEA)

Dalam setiap Bab, disusun dengan sistematika sebagai berikut :

A. PENDAHULUAN

B. PENYAJIAN MATERI

C. RANGKUMAN

D. TUGAS

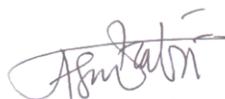
E. RUJUKAN

F. BACAAN YANG DIANJURKAN

Seperti halnya bunyi pepatah, “*tiada gading yang tak retak*”, maka penulis menyadari masih banyak yang perlu disempurnakan dalam BUKU AJAR ini dan saran-saran dari semua pihak sangat diharapkan untuk penyempurnaannya.

Akhirnya kami mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang membantu penyelesaian Buku Ajar ini. Mudah-mudahan Buku Ajar ini bermanfaat bagi kita semua.

Pamekasan, Juli 2021



**Penulis**

## Sanwacana

Puji syukur kehadiran Allah SWT. yang telah menganugerahkan begitu banyak rahmat, karunia, dan hidayah-Nya, sehingga penulis dapat menyelesaikan Buku Ajar dengan judul “STATISTIK BISNIS (TEORI DAN APLIKASI PROGRAM R (Paket R-Commander). *Sholatullah wasalamuhu* semoga senantiasa tercurahkan kepada Nabi penyempurna akhlak manusia, Baginda Muhammad SAW. Buku Ajar ini merupakan hasil penyusunan sebagai salah satu syarat untuk memperoleh jabatan Tenaga Edukatif Dosen di IAIN MADURA.

Penyusunan Buku Ajar ini tidak lepas dari bantuan berbagai pihak, oleh karena itu penulis ingin menyampaikan ucapan terima kasih kepada:

1. Dr. H. Mohammad Kosim, M.Ag selaku rektor dan jajarannya di rektorat khususnya Dr H. Nor Hasan, M.Ag selaku wakil rektor bidang Akademik dan Pengembangan Kelembagaan yang telah membuat program tentang buku aja ini
2. Tim LPM IAIN MADURA yang telah memfasilitasi program penulisan buku ajar ini mulai dari mengadakan peatihan sampai terbitnya buku ajar ini
3. Dr. H. Zainal Abidin, MEI selaku dekan Fakultas Ekonomi dan Bisnis Islam dan jajarannya khususnya wadek bidang bidang Akademik dan Kemahasiswaan Bapak Dr. H. Rudy Haryanto, SST. MM yang telah banyak mengarahkan sehingga penulis dapat ditentukan judul buku ajar yang sesuai bidang keahlian dan Bapak H. Wadhan, SE. M.Si selaku Kepala Laboratorium FEBI yang telah memfasilitasi Lab sebagai tempat kami berkumpul dalam diskusi menulis Buku Ajar.
4. Ibu Fatati Nuryana, M.Si dan Khotibul Umam, M.E.I. selaku rivewer yang telah banyak memberikan bimbingan dan pengarahan sehingga Buku Ajar ini terselesaikan dengan baik;

5. Sahabat-sahabatku tercinta, khususnya CPNS di Fakultas Ekonomi dan Bisnis Islam dan seluruh CPNS ATM 2019 yang tidak bisa saya sebutkan satu-persatu, terima kasih atas semangat dan motivasinya selama penulisan buku ajar ini
6. Istriku faizatul Abadiyah dan anak-anakku Mohammad Ahsanul Haq dan Alisyah Almira Ramadhani terimakasih atas pengertian kerelaan tersitanya waktu kebersamaan kita karena banyak terlupakan untuk menyelesaikan penulisan Buku Ajar ini.

Akhirnya penulis berharap dan memohon semoga bantuan semua pihak mendapat ridho dan bernilai amal ibadah yang saleh serta mendapat balasan yang setimpal dari Allah SWT dan semoga Buku Ajar ini dapat memberikan manfaat bagi penulis khususnya dan bagi pembaca pada umumnya. *Aamiin...*

Pamekasan, Juli 2021



**Penulis**

# DAFTAR ISI

<b>HALAMAN PENGESAHAN</b> .....	<b>iii</b>
<b>Kata Sambutan</b> .....	<b>iv</b>
<b>Kata Pengantar</b> .....	<b>v</b>
<b>Prakata</b> .....	<b>vi</b>
<b>Sanwacana</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>x</b>
<b>BAB I</b>	
<b>KONSEP DASAR STATISTIKA</b> .....	<b>1</b>
A. Pendahuluan .....	1
B. Peranan Statistik .....	2
C. Rangkuman.....	13
D. Tugas .....	14
E. Rujukan .....	14
F. Bacaan yang Dianjurkan.....	15
<b>BAB II</b>	
<b>PROGRAM STATISTIKA R</b> .....	<b>16</b>
A. Pendahuluan .....	16
B. Pengenalan Program R.....	16
C. Rangkuman.....	25
D. Tugas .....	26
E. Rujukan .....	27
F. Bacaan yang Dianjurkan.....	27
<b>BAB III</b>	
<b>STATISTIKA DESKRIPTIF</b> .....	<b>29</b>
A. Pendahuluan .....	29
B. Penyajian Data.....	29
C. Rangkuman.....	49

D. Tugas .....	50
E. Rujukan .....	50
F. Bacaan yang Dianjurkan .....	50
<b>BAB IV</b>	
<b>PENGUJIAN HIPOTESIS.....</b>	<b>51</b>
A. Pendahuluan .....	51
B. Konsep Uji Hipotesis .....	51
C. Rangkuman.....	58
D. Tugas .....	59
E. Rujukan .....	60
F. Bacaan yang Dianjurkan .....	61
<b>BAB V</b>	
<b>PENGUJIAN DATA.....</b>	<b>62</b>
A. Pendahuluan .....	62
B. Konsep Uji Normalitas Data .....	62
C. Rangkuman.....	66
D. Tugas .....	67
E. Rujukan .....	67
F. Bacaan yang Dianjurkan .....	68
<b>BAB VI</b>	
<b>UJI BEDA RATA-RATA.....</b>	<b>69</b>
A. Pendahuluan .....	69
B. Pengujian Beda Rata-Rata.....	70
C. Rangkuman.....	85
D. Tugas .....	87
E. Rujukan .....	87
F. Bacaan yang Dianjurkan .....	88

## **BAB VII**

### **PENGUJIAN HIPOTESIS TENTANG PROPORSI ..... 89**

- A. Pendahuluan .....89
- B. Uji Proporsi.....90
- C. Rangkuman.....98
- D. Tugas ..... 100
- E. Rujukan ..... 100
- F. Bacaan yang Dianjurkan..... 101

## **BAB VIII**

### **TABEL KONTINGENSI .....102**

- A. Pendahuluan ..... 102
- B. Konsep Tabel Kontingensi ..... 102
- C. Rangkuman..... 111
- D. Tugas ..... 114
- E. Rujukan ..... 114
- F. Bacaan yang Dianjurkan..... 114

## **BAB IX**

### **PENGUJIAN INSTRUMEN PENELITIAN ..... 115**

- A. Pendahuluan ..... 115
- B. Konsep Pengukuran Validitas dan Reliabilitas..... 116
- C. Rangkuman..... 144
- D. Tugas ..... 149
- E. Rujukan ..... 149
- F. Bacaan yang Dianjurkan..... 149

## **BAB X**

### **STATISTIKA NON PARAMETRIK ..... 151**

- A. Pendahuluan ..... 151
- B. Konsep Statistika Non Parametrik ..... 152
- C. Rangkuman..... 170

D. Tugas .....	172
E. Rujukan .....	172
F. Bacaan yang Dianjurkan .....	173
<b>BAB XI</b>	
<b>ANALISIS KORELASI .....</b>	<b>174</b>
A. Pendahuluan .....	174
B. Konsep Analisis Korelasi .....	175
C. Rangkuman .....	187
D. Tugas .....	188
E. Rujukan .....	188
F. Bacaan yang Dianjurkan .....	188
<b>BAB XII</b>	
<b>ANALISIS REGRESI .....</b>	<b>189</b>
A. Pendahuluan .....	189
B. Konsep Analisis Regresi .....	189
C. Rangkuman .....	210
D. Tugas .....	211
E. Rujukan .....	212
F. Bacaan yang Dianjurkan .....	212
<b>BAB XIII</b>	
<b>METODE DATA ENVELOPMENT ANALYSIS (DEA) .....</b>	<b>214</b>
A. Pendahuluan .....	214
B. Teori Metode Data Envelopment Analysis (Dea) .....	215
C. Rangkuman .....	229
D. Tugas .....	230
E. Rujukan .....	230
F. Bacaan yang Dianjurkan .....	230

<b>DAFTAR PUSTAKA .....</b>	<b>232</b>
<b>GLOSARI .....</b>	<b>236</b>
<b>BIOGRAFI PENULIS .....</b>	<b>242</b>

# BAB I

## KONSEP DASAR STATISTIKA

### A. Pendahuluan

Pada Bab ini akan dibahas hal-hal yang mendasar berkenaan dengan statistika dimulai dari peranan statistik, sumber dan jenis data, tingkat skala pengukuran dan peran penggunaan komputer yang begitu penting dalam statistika.

#### **Tujuan Pembelajaran :**

Setelah mempelajari materi pada bab ini mahasiswa diharapkan memahami prinsip konsep dasar statistika, peranan statistika dalam persoalan ekonomi dan bisnis.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nilai  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Peranan Statistik.
2. Sumber Dan Jenis Data.
3. Tingkat Skala Pengukuran Data.
4. Penggunaan Komputer.

## **B. Peranan Statistik**

Sebagai suatu disiplin ilmu, Statistik pada mulanya berkembang karena kebutuhan pihak pemerintah atau penguasa untuk mengumpulkan informasi yang berkaitan dengan data kependudukan. Sejarah menunjukkan bahwa pengumpulan data telah dilakukan sejak berabad-abad yang lalu. Pada masa kejayaan kekaisaran Yunani dan Romawi pengumpulan informasi dilakukan terutama sekali untuk keperluan penentuan pajak dan ketentuan yang berkaitan dengan wajib dinas militer. Demikian juga pemerintah Belanda ada pada masa penjajahannya di Indonesia telah mengumpulkan berbagai informasi tentang Indonesia, tidak hanya tentang kependudukan saja tetapi juga berbagai informasi lainnya, seperti informasi tentang flora dan fauna Indonesia, Sosial, budaya, dan sebagainya. Saat ini semua negara di dunia telah melakukan berbagai jenis sensus secara periodik, paling tidak setiap 10 tahunan, mendapatkan berbagai informasi yang akurat dan up to date. Istilah-istilah seperti sensus kependudukan, sensus pertanian, dan sensus ekonomi, sudah bukan merupakan istilah yang asing lagi bagi telinga kita.

Terdapat dua istilah yang penting yang sering dikaitkan dengan statistik yaitu populasi dan sampel. dengan penjelasan populasi adalah seluruh nilai atau item yang mungkin diperoleh sebagai hasil pengamatan dalam suatu persoalan tertentu. sedangkan sampel adalah bagian dari populasi yang diambil untuk keperluan analisis.

Berbeda dengan pengertian sehari-hari, istilah populasi dalam statistik tidak harus selalu dikaitkan dengan sekelompok orang tertentu. Populasi bisa saja berkaitan dengan diameter pohon suatu spesies pada umur tertentu, atau berat kering tanaman dari suatu varietas cabai setelah dipanen, atau penghasilan perbulan dari golongan masyarakat

tertentu. Ukuran deskriptif dari suatu populasi disebut parameter. misalnya untuk kasus di mana populasinya adalah penghasilan perbulan, maka parameter populasi tersebut dapat berupa rata-rata penghasilan perbulan atau proporsi golongan masyarakat tersebut yang berpenghasilan di bawah suatu nilai standar tertentu. Jika tak hanya mengamati penghasilan perbulan dari 100 orang dalam golongan masyarakat tadi, maka nilai-nilai yang kita peroleh merupakan sampel dari populasi tersebut. Ukuran deskriptif dari suatu sampel disebut statistik.

Istilah statistik saat ini mempunyai dua arti yang berbeda. pada umumnya statistik berarti Sekumpulan data yang terdiri atas angka angka, sehingga kita mengenal statistik pertanian, statistik perekonomian, statistik penduduk, dan sebagainya. Selain itu, statistik juga berarti suatu disiplin ilmu. ruang lingkup statistik sebagai disiplin ilmu mencakup berbagai teknik pengumpulan dan penyajian data, baik untuk keperluan analisis data maupun dalam proses pengambilan keputusan. Salah satu kegunaan utama dari ilmu statistik adalah untuk menyediakan suatu prosedur yang memungkinkan kita melakukan inferensi pendugaan dan menentukan keputusan tentang karakteristik suatu populasi berdasarkan atas informasi sampel yang diambil dari sebagian populasi tersebut.

Ilmu Statistik dapat dibagi menjadi dua bagian, yaitu statistik deskriptif dan statistik inferensial. Cabang ilmu statistik deskriptif berkaitan dengan berbagai teknik pengumpulan organisasi yang penyederhanaan dan penyajian data ke dalam bentuk yang lebih mudah dipahami misalnya dalam bentuk tabel atau grafik. Teknik penyederhanaan data biasanya disertai dengan penjelasan tentang karakteristik karakteristik tertentu dari data seperti ukuran pemusatan data dan ukuran penyebaran data.

Saat ini statistik telah digunakan secara luas dalam penelitian Penelitian di semua disiplin ilmu pengetahuan. hal ini terutama sekali terjadi setelah berkembangnya metode statistik inferensial yang merupakan perkembangan dari teori peluang. statistik inferensial berkaitan dengan proses pendugaan dan penarikan kesimpulan tentang karakteristik suatu populasi berdasarkan atas informasi sampel.

Salah satu aspek utama dari statistik inferensial adalah proses penggunaan nilai statistik sampel dalam pengambilan keputusan berkaitan dengan nilai parameter populasi yang sebenarnya. dengan semakin besarnya populasi maka biaya ya dan waktu yang diperlukan untuk mendapatkan informasi dari seluruh populasi akan semakin besar dan semakin sulit dilaksanakan. sehingga kesimpulan mengenai karakteristik populasi harus ditentukan berdasarkan informasi yang diperoleh dari sampel yang diambil dari populasi tersebut. teori peluang mempunyai peranan yang besar dalam menjembatani hubungan antara hasil sampel dengan populasinya. kaidah-kaidah dalam teori peluang digunakan untuk menilai sampai Berapa jauh informasi sampel mencerminkan populasinya. penilaian tersebut dilakukan melalui pemeriksaan terhadap sifat-sifat distribusi sampling nya.

Aplikasi statistika dalam berbagai bidang contohnya dalam bidang SosPoL (Sosiometrika) yaitu Polling masyarakat, Pengambilan Keputusan, Klasifikasi potensi di Perusahaan/Pabrik melakukan Kontrol Kualitas, Pengambilan Keputusan. Di bidang Biosains (Biometrika/BioInfortmika) yaitu Menguji hasil eksperimen (obat dan produk lainnya), Rancob, Regresi. Dibidang Psikologi dan Pendidikan (Psikometrika) yaitu Mengukur dan menguji hubungan berbagai kemampuan dan karakter orang, termasuk validitas alat ukur (SEM). Dibidang Ekonomi (Ekonometrika) yaitu

Mempelajari hubungan antar peubah-peubah ekonomi, trend terhadap waktu (*Time Series*).

Aplikasi statistika dalam bidang politik dan keamanan yaitu ditunjukkan dengan berbagai media yang secara periodik mengadakan jajak pendapat tentang penilaian masyarakat terhadap suatu kebijakan pemerintah maupun penilaian mereka tentang kemungkinan ketua-ketua partai besar untuk memenangkan pemilihan. Hasil jajak pendapat umumnya dinyatakan dalam angka prosentase setuju atau tidak setuju, percaya atau tidak percaya, maupun prosentase perolehan suara dan sebagainya. Demikian juga setiap akhir tahun, kepolisian memberikan laporan tentang kenaikan atau penurunan angka kejahatan. Baik disuatu wilayah tertentu maupun secara nasional. Semua ini merupakan sebagian dari kegiatan statistika dalam bidang politik dan keamanan.

Aplikasi statistik dalam Bidang kesehatan dan pertanian Dokter peneliti atau insinyur pertanian mengadakan percobaan untuk menentukan efek dari berbagai obat-obatan dan mengontrol kondisi lingkungan pada manusia untuk memutuskan pengobatan yang tepat untuk berbagai penyakit. Demikian juga efektifitas dari penggunaan makanan atau obat-obatan suplemen baik untuk manusia maupun tanaman dalam bidang pertanian. Dalam bidang manufaktur secara internasional peranan statistika dalam mengontrol kualitas produksi ditunjukkan oleh Negara Jepang. Misalnya pabrik mobil Toyota sangat sungguh-sungguh dalam mengumpulkan dan menganalisis data tentang kualitas produksi berikutnya. Secara umum, dalam bidang manufaktur, para peneliti mengambil sampel karakteristik kualitas suatu produk dan berbagai peubah yang dapat dikontrol untuk mengidentifikasi variable kunci yang berhubungan dengan kualitas produk.

Aplikasi statistika Dalam bidang bisnis dan ekonomi, dalam bidang ini statistika, misalnya dipegunakan untuk mengambil sampel pelanggan untuk memperoleh informasi guna memprediksi kesukaan terhadap suatu produk. Barang yang baru diproduksi biasanya disampel sebelum didistribusikan untuk menentukan apakah memenuhi syarat atau tidak. Demikian juga penentuan jaminan purna jual tidak lepas dari hasil pengujian beberapa produk sabagai sampel. Para ekonom mengamati berbagai indeks kesehatan ekonomi selama beberapa periode waktu dan menggunakan informasi yang diperoleh untuk meramalkan kondisi ekonomi di masa depan. Media-media setiap hari melaporkan harga rata-rata kebutuhan pokok. Badan Pusat Statistika misalnya, secara periodic melaporkan angka pengngguran dan inflasi. Contoh-contoh tersebut menunjukkan sebagian pemanfaatan statistika dalam berbagai bidang diatas.

### **1. Sumber Dan Jenis Data**

Informasi atau data selalu diperlukan dalam setiap proses pengambilan keputusan. Akan tetapi, tidak semua data Memberikan manfaat dalam pengambilan keputusan. jika data yang diperoleh tidak layak atau cacat karena bias, tidak jelas, as atau karena kesalahan kesalahan lainnya, maka tidak ada satupun alat atau metode yang dapat memperbaikinya. oleh karena itu, metode pengumpulan data yang baik dan benar perlu mendapat perhatian yang serius, agar data yang diperoleh Memberikan manfaat yang maksimal.

Terdapat beberapa metode yang dapat digunakan untuk mendapatkan data yang diperlukan, yang mencari data yang sudah dipublikasikan oleh sumber-sumber tertentu baik pemerintah perusahaan atau individu. merancang percobaan. dan melakukan survey.

Pemerintah di setiap tingkatan, baik pusat, propinsi maupun kabupaten, telah secara rutin mempublikasikan berbagai jenis data melalui badan pusat statistic di masing-masing tingkat pemerintahan. Data yang dipublikasikan pemerintah diantaranya mencakup data kependudukan, tenaga kerja, pertanian, perekonomian, dan sebagainya. Data tersebut digunakan pemerintah sebagai pijakan untuk mengambil keputusan atau kebijakan dan program-program pembangunan diberbagai bidang. Selain itu, Berbagai jenis data Juga dikumpulkan oleh perusahaan, lsm, dan peneliti dari berbagai bidang. data yang dikumpulkan oleh pihak nonpemerintah biasanya hanya digunakan untuk keperluan sendiri dalam lingkungan yang terbatas dan tidak selalu dipublikasikan oleh umum. Namun demikian data tersebut seringkali dapat diakses oleh pihak lain untuk berbagai keperluan lain. data yang demikian disebut sebagai data sekunder. karena data tersebut telah dikompilasi dan telah siap untuk digunakan untuk keperluan analisis selanjutnya.

Data dapat juga dikumpulkan melalui percobaan atau eksperimen. di dalam suatu percobaan, pengaturan dan kontrol yang ketat diterapkan terhadap satuan-satuan percobaan yang mendapat perlakuan tertentu. misalnya, dalam suatu percobaan untuk menguji pertumbuhan beberapa varietas padi pada suatu jenis tanah tertentu, maka keadaan tanah, pengairan dan faktor-faktor lain yang dapat mempengaruhi pertumbuhan tanaman padi harus dibuat seragam mungkin. hal ini dimaksudkan agar perbedaan pertumbuhan antara tanaman yang dihasilkan melalui percobaan tersebut diharapkan mencerminkan perbedaan varietas.

Metode yang ketiga dalam pengumpulan data adalah melalui survey. dalam pelaksanaan suatu survei, pengaturan dan kontrol terhadap satuan percobaan atau objek yang disurvei tidak diterapkan secara ketat seperti Pada pelaksanaan suatu percobaan. pengumpulan data melalui survei biasanya dilakukan melalui teknik wawancara, penyebab, di diskusi atau pengamat objek di teliti.

Data yang dikumpulkan sendiri oleh peneliti, baik melalui suatu percobaan maupun melalui survei, di sebut dengan data primer. Data yang demikian dikumpulkan secara khusus agar sesuai dengan keperluan analisis yang diinginkan.

Pada dasarnya data yang diperoleh melalui ketiga metode tersebut dapat dikelompokkan menjadi dua jenis data, yaitu data kualitatif dan data kuantitatif. data kualitatif adalah ah da tayang yang berbentuk kategori. misalnya data tentang jenis kelamin, warna dan jenis pekerjaan. sedangkan data kuantitatif adalah ah ah da ta yang berbentuk numerik atau angka, misalnya data tentang hasil tanaman per hektar, penghasil, dan berat badan. data kuantitatif dapat dikelompokkan lebih jauh menjadi data kuantitatif diskrit dan kontinu. data yang data yang biasa membilang, superti jumlah buah per tanaman, atau jumlah anggota keluarga. data yang diskrit hanya dapat mengambil nilai pada suatu titik tertentu dalam suatu selang atau interval, sehingga selalu ada jarak atau celah di antara nilai-nilainya. oleh karena itu, data kuantitatif yang di script biasanya mempunyai nilai berupa bilangan bulat misalnya 0,1,2 dan seterusnya. sebaliknya data yang kontinyu biasanya diperoleh sebagai hasil pengukuran, dan nilainya dapat mengambil sembarang nilai dalam suatu interval tertentu, misalnya

berat buah, tinggi tanaman, atau penghasil perbulan. nilai pengamatan dari data kuantitatif yang kontinu sangat tergantung pada tingkat ketelitian alat ukurnya. misalnya tanaman yang dilaporkan tingginya adalah 16meter jika diukur lebih teliti mungkin hasil pengukuran 16,2 m. oleh karena itu, data kuantitatif data kuantitatif kontinu dapat mengambil sembarang nilai dalam sistem bilangan nyata. dalam suatu survei, jenis data yang ingin dikumpulkan dapat disesuaikan dengan bentuk pertanyaan yang diajukan kepada responden serta responden dari pertanyaan tersebut.

## **2. Tingkat Skala Pengukuran Data**

Pada umumnya Data diperoleh dengan cara melakukan pengukuran atau penilaian terhadap objek yang diamati. Oleh karena itu data dapat juga digolongkan berdasarkan tingkat skala pengukurannya. yaitu skala nominal, skale ordinal, skala interval, dan skala rasio. Makin tinggi tingkat skala pengukurannya, makin banyak informasi yang dikandung data tersebut.

Skala pengukuran nominal dan ordinal hanya terdapat dalam data kualitatif. data yang diukur pada skala nominal merupakan tingkatan terendah dalam skala pengukuran. dalam skala nominal data hanya dikelompokkan kedalam beberapa kategori yang berbeda tanpa adanya tingkatan. setiap objek pengamatan hanya dapat dikelompokkan kedalam satu kategori saja. beberapa contoh data dalam skala pengukuran nominal yaitu kepemilikan roda dua, jenis kewarganegaraan, dan warna daun. jika pengelompokan data kedalam beberapa kategori disertai dengan adanya tingkatan dalam kategori tersebut maka yang terjadi adalah pengukuran dalam skala ordinal. data yang diukur

dalam skala ordinal mempunyai tingkatan yang lebih tinggi daripada data yang diukur dalam skala nominal. Namun demikian, skala pengukuran ordinal masih merupakan skala pengukuran yang lemah, karena kita tidak bisa membuat perbedaan kategori tersebut. perbedaan kategori tersebut hanya sebatas Pada kategori mana yang lebih baik atau lebih tinggi atau lebih disukai dan dan tidak lebih dari itu. contohnya jabatan dosen guru besar lektor kepala lektor dan asisten ahli, tingkat hotel berbintang.

Skala pengukuran interval dan rasio berlaku bagi data kuantitatif. skala pengukuran interval diperoleh ketika objek pengamatan diukur secara numerik dan interval antara hasil hasil pengukuran dapat ditentukan dengan tepat. pola interval merupakan skala pengukuran dalam bentuk yang lebih tegas daripada skala nominal dan ordinal. dalam skala interval perbedaan antara hasil pengukuran menghasilkan suatu besaran yang konstan. misalnya, pada pengukuran larutan lain yang suhunya 48 derajat Celcius. Namun demikian, data yang diukur dengan skala interval mempunyai titik nol sembarang atau berubah-ubah. misalnya, titik nol pada pengukuran suhu dengan skala Fahrenheit berbeda dengan titik nol pada skala Celcius. hal ini berbeda dengan data yang diukur dalam skala pengukuran rasio. data yang diukur pada skala rasio selain mempunyai interval pengukuran yang tepat juga mempunyai titik nol yang tetap. salah satu contoh pengukuran dalam skala rasio adalah pada pengukuran umur bola lampu pijar. perbedaan umur antara bola lampu yang berumur 1000 jam dengan yang berumur 500jam. bola lampu pijar yang berumur 1000 jam dikatakan dua kali lebih tahan dari pada bola lampu pijar

yang berumur 500 jam. Selain itu pengukuran ini mempunyai titik nol yang tetap bola lampu yang berumur 0 jam berarti bola lampu yang sama sekali tidak menyala. hal ini berbeda dengan pengukuran suhu larutan, kita tidak bisa mengatakan bahwa larutan yang mempunyai suhu 102 kali lebih panas daripada larutan yang mempunyai suhu 50 derajat Celcius. contoh lain untuk skala pengukuran interval yaitu penanggalan kalender (Masehi, Hijriah, Cina atau Saka). contoh lain untuk skala pengukuran rasio tinggi badan, berat badan dan umur.

Metode statistik yang digunakan untuk menganalisis data kualitatif berbeda dengan metode yang digunakan untuk data kuantitatif. oleh karena itu pengetahuan tentang jenis dan skala pengukuran data sangatlah penting, karena penggunaan metode yang tidak tepat dapat menyebabkan hasil analisis yang salah.

### **3. Penggunaan Komputer**

Ukuran data semakin besar dan kompleks sehingga semakin sulit/ mustahil dikerjakan secara manual. Banyak teknik analisis statistika yang sulit dilakukan secara manual, tetapi sudah banyak terimplementasi di computer. Memudahkan ilustrasi dan eksplorasi grafik, Memungkinkan pemeriksaan asumsi dengan cepat, Memungkinkan mencoba berbagai alternatif dengan cepat dan memilih yang terbaik. Macam-macam software statistika diantaranya R merupakan Identik dengan SPlus (kelebihannya gratis, untuk dasar sudah berbasis menu, berbahasa Indonesia). SPSS Cocok untuk bidang sosial (multivariate). MINITAB Cocok untuk bidang biosains (Rancob). S-Plus Cocok untuk bidang biosains (Modeling, Regresi). SAS Paket Sangat Komprehensif, tapi mahal.

Dengan semakin meluasnya ketersediaan program pada komputer di lembaga pendidikan akhir-akhir ini, metode pengajaran statistik di beberapa perguruan tinggi telah mengalami perubahan yang drastis. Demikian juga penggunaan metode statistik pada penelitian di berbagai disiplin ilmu telah mengalami kemajuan yang pesat. Program-program statistik yang telah banyak diciptakan dan selalu diperbarui untuk mengikuti perkembangan di bidang ilmu statistik dan untuk kemudahan bagi penggunaannya. Dengan menggunakan program-program paket statistik analisis dan manipulasi data menjadi jauh lebih mudah, bahkan untuk data yang besar sekalipun. Perhitungan-perhitungan statistik yang dulu sangat rumit dan sangat menyita waktu kini dapat dilakukan dengan mudah dan dalam waktu yang sangat singkat hanya dengan Menuliskan beberapa perintah pemrograman atau bahkan cukup dengan menggunakan menu yang sudah tersedia. Dalam buku ini penggunaan program paket statistik yang digunakan yaitu bersifat open source artinya bebas dalam memilikinya tidak perlu berlisensi yaitu program statistik R. Dalam program ini sebenarnya relatif rumit akan tetapi sudah terdapat aplikasi paket yang berbentuk pull down menu yaitu paket R Commander. Artinya menu yang tertera pada menu bar terdiri atas submenu yang dapat diaktifkan ketika menu tersebut diklik dengan mouse komputer. Selain itu penggunaan program R ini merupakan salah satu program yang sering digunakan oleh para ahli statistik di dunia. Karena bersifat ekonomis tidak perlu mengeluarkan biaya untuk memilikinya dengan membayar lisensi. Dan program R ini merupakan suatu program statistik yang sangat canggih dan terus

berkembang secara bebas apabila ingin menekuni secara khusus pengembangan paket yang ada dalam program R.

### C. Rangkuman

Populasi adalah seluruh nilai atau item yang mungkin diperoleh sebagai hasil pengamatan dalam suatu persoalan tertentu. Sampel adalah bagian dari populasi yang diambil untuk keperluan analisis. Ukuran deskriptif dari suatu sampel disebut statistik. Ilmu statistik adalah untuk menyediakan suatu prosedur yang memungkinkan kita melakukan inferensi pendugaan dan menentukan keputusan tentang karakteristik suatu populasi berdasarkan atas informasi sampel yang diambil dari sebagian populasi tersebut.

Data dapat juga dikumpulkan melalui percobaan atau eksperimen. Data yang dikumpulkan sendiri oleh peneliti, baik melalui suatu percobaan maupun melalui survei, disebut dengan data primer. Data yang dihimpun oleh pihak lain adalah data sekunder. Skala pengukuran nominal dan ordinal hanya terdapat dalam data kualitatif. Skala pengukuran interval dan rasio berlaku bagi data kuantitatif.

Ukuran data semakin besar dan kompleks sehingga semakin sulit/mustahil dikerjakan secara manual. Banyak teknik analisis statistika yang sulit dilakukan secara manual, tetapi sudah banyak terimplementasi di computer. Memudahkan ilustrasi dan eksplorasi grafik, Memungkinkan pemeriksaan asumsi dengan cepat, Memungkinkan mencoba berbagai alternatif dengan cepat dan memilih yang terbaik. Macam-macam software statistika diantaranya R merupakan Identik dengan SPlus (kelebihannya gratis, untuk dasar sudah berbasis menu, berbahasa Indonesia). SPSS Cocok untuk bidang sosial (multivariate). MINITAB Cocok untuk bidang biosains (Rancob). S-Plus Cocok untuk bidang biosains

(Modeling, Regresi). SAS Paket Sangat Komprehensif, tapi mahal.

#### **D. Tugas**

1. Jelaskan perbedaan antara populasi dan sampel ?
2. Jelaskan apa yang dimaksud dengan parameter ?
3. Sebutkan dua cabang ilmu statistic dan jelaskan ruang lingkupnya masing-masing?
4. Jelaskan peranan statistic sebagai ilmu bagi dunia ilmu pengetahuan?
5. Jelaskan antara data primer dan data sekunder
6. Berikan lima contoh populasi dan sampelnya serta ukuran deskriptif masing-masing
7. Sebutkan tiga metode yang biasa digunakan untuk mendapatkan data dan berikan contohnya
8. Jelaskan perbedaan antara skala data intervan, nominal, rasio dan ordinal, berikan lima contoh dari masing-masing skala pengukuran
9. Jelaskan perbedaan antara data diskrit dan data kontinyu, berikan lima contoh dari masing-masing jenis data tersebut
10. Jelaskan perbedaan antara data kualitatif dan data kuantitatif, berikan lima contoh dari masing-masing jenis data tersebut

#### **E. Rujukan**

Furqon, (2001) *Statistika Terapan untuk Penelitian*, Bandung: Alfabeta

Hadi, S (1995), *Statistik 1, 2, 3*, Yogyakarta: Andi Offset

Nuryadi, dkk. (2017). *Dasar-dasar Statistik Penelitian*. Yogyakarta: Gramasurya.

Sudjana, (1992), *Metoda Statistika (Edisi ke 5)*, Bandung: Tarsito

Sugiyono (2003), *Statistika untuk Penelitian*, Bandung: Alfabeta

Supramono dan Sugiarto (1993), *Statistika*, Yogyakarta: Andi Offset.

#### **F. Bacaan yang Dianjurkan**

Furqon, (2001) *Statistika Terapan untuk Penelitian*, Bandung: Alfabeta

Hadi, S (1995), *Statistik 1, 2, 3*, Yogyakarta: Andi Offset

Nuryadi, dkk. (2017). *Dasar-dasar Statistik Penelitian*. Yogyakarta: Gramasurya.

Sudjana, (1992), *Metoda Statistika (Edisi ke 5)*, Bandung: Tarsito

Sugiyono (2003), *Statistika untuk Penelitian*, Bandung: Alfabeta

Supramono dan Sugiarto (1993), *Statistika*, Yogyakarta: Andi Offset.

## BAB II

# PROGRAM STATISTIKA R

### A. Pendahuluan

Pada Bab ini akan dibahas tentang Program Statistika R, mengaktifkan R-Commander, Mengentri Data, dan Impor Data beserta prakteknya dalam persoalan ekonomi dan bisnis.

#### **Tujuan Pembelajaran :**

Setelah mempelajari materi pada bab ini mahasiswa diharapkan mampu mengaplikasikan Program Statistika R, mengaktifkan R-Commander, Mengentri Data, dan Impor Data beserta prakteknya dalam persoalan ekonomi dan bisnis.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nila  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Mengaplikasikan Program Statistika R
2. Mengaplikasikan Mengaktifkan R-Commander
3. Mengaplikasikan Entri Data
4. Mengaplikasikan Impor Data

### B. Pengenalan Program R

R adalah salah satu *software* analisis data yang termasuk kelompok *software* statistik yang *open source* yang

tidak memerlukan lisensi atau gratis, yang dikenal dengan *freeware*. R dapat diperoleh secara gratis di <https://cran.r-project.org>. *software* R ini sudah Dilengkapi dengan banyak kemampuan internal untuk menganalisis data dan menampilkan grafik sehingga *software* R bisa dikategorikan sebagai paket pengolahan data (*software* statistika).

Beberapa kemampuan yang menonjol dari *software* R yang menjadi alasan banyak statistisi memilihnya sebagai program aplikasi antara lain yaitu, program R memiliki koleksi program analisis data yang disebut program R memiliki koleksi program analisis data, yang disebut library atau pustaka yang sangat luas seperti statistika deskriptif, regresi, pemodelan statistika (baik linier maupun non linier), Anova dan multivariat. variasi penampilan grafiknya sangat banyak dan berkualitas tinggi, baik penampilan di layar monitor maupun dalam bentuk cetak di atas. Kemampuan pemrograman (bahasa S) dapat dikembangkan secara fleksibel untuk kepentingan khusus yang lebih lanjut. Merupakan pemrograman yang berorientasi pada objek. keuntungannya, apabila apa yang telah dikerjakan R saat ini diperlukan di kemudian hari maka R dapat mengambilnya tanpa harus melakukan perhitungan ulang dari awal.

### **1. Mengaktifkan Program R**

Apabila program R telah ter-install dengan baik pada komputer anda. Untuk mengaktifkan program R tersebut, langkah-langkahnya yaitu sebagai berikut, klik dua kali pada shortcut R di desktop windows seperti pada gambar berikut :



Program R yang tersedia paling terbaru versi **R-4.1.0 for Windows (32/64 bit)**. Tampilan awal program R akan terbuka seperti yang terlihat pada gambar berikut

A screenshot of the R Console (32-bit) window. The window title bar reads "R Console (32-bit)". The menu bar includes "File", "Edit", "Misc", "Packages", "Windows", and "Help". The main content area displays the following text:

```
R version 4.1.0 (2021-05-18) -- "Camp Pontanezen"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/i386 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

## 2. Mengaktifkan R-Comannder

Dalam program R sudah terdapat aplikasi paket berbasis menu disebut dengan R Commander. R Commander adalah salah satu R-Gui (tampilan program R) yang dapat digunakan untuk berbagai analisis statistika dasar. Paket R-Commander dapat dijalankan dengan dua cara yang berbeda yaitu dengan mengetikkan perintah `library(Rcmdr)` pada tampilan atau disebut jendela R-console dan menekan enter

```
R Console (32-bit)
File Edit Misc Packages Windows Help

R version 4.1.0 (2021-05-18) -- "Camp Pontanezen"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/i386 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

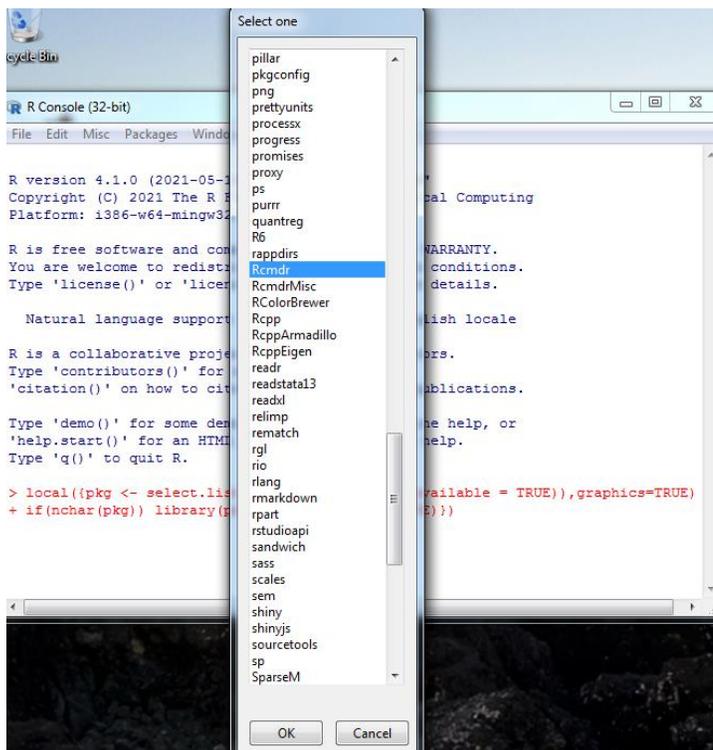
Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

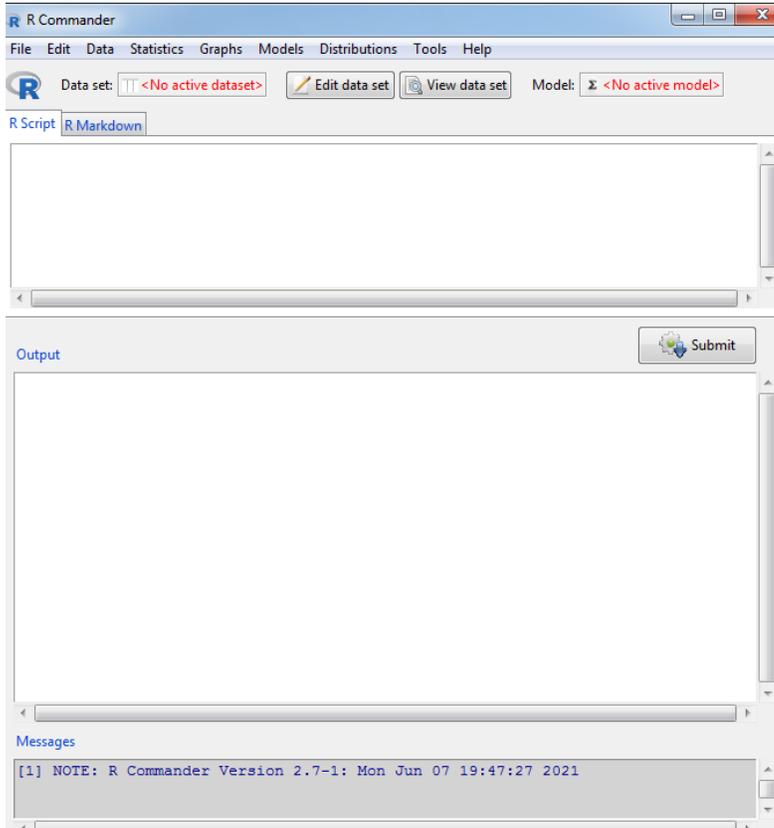
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> library(Rcmdr)|
```

atau mengklik menu Packages, dan pilih Load packages dan kemudian pilih Rcmdr pada daftar paket library yang telah terinstall, seperti pada gambar berikut



Paket R-Comannder akan terbuka seperti pada gambar berikut



### 3. Mengentri Data

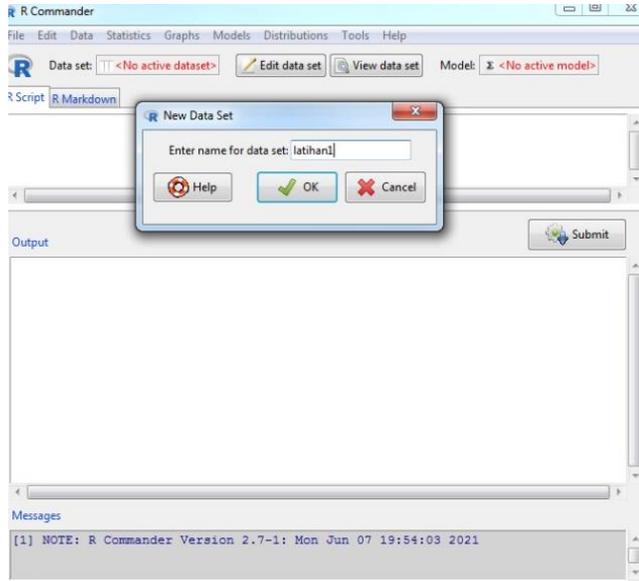
Secara umum proses input data di R-commander dapat dilakukn dengan dua macam cara yaitu data diinputkan langsung melalui data editor, dan data yang diberikan oleh program lain melalui impor data. Program yang format datanya dapat dibaca oleh program R yaitu data dari file txt atau clipboard, dataset SPSS, dataset MINITAB, dataset STATA, data dari excel, Access atau dBase.

#### a. Data editor

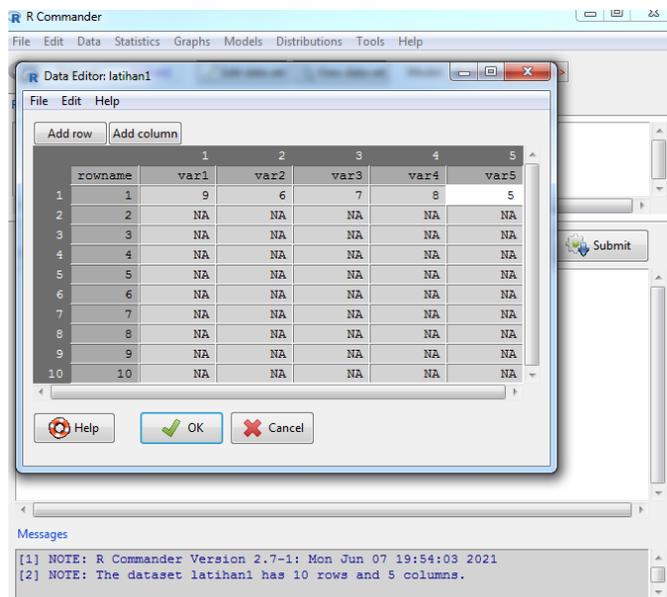
Adapun cara input data secara langsung melalui data editor dengan neggunakan R-Commander yaitu

Klik menu Data-> Data Set baru

Tampilan dialog pengisian nama data set akan muncul, masukkan nama untuk data set (misalkan latihan 1) kemudian klik tombol ok.



Tampilan data editor untuk pengisian data akan ditampilkan seperti pada gambar berikut



Untuk memberikan nama variable yaitu dengan cara mengklik pada kolom paling atas dari data editor (var1, var2, var3, dst).

Tampilan variable editor untuk pengisian nama variable akan muncul. Tuliskan nama variabelnya dan pilih type data yang akan diisi, dengan cara numeric jika datanya berupa angka dan character jika datanya berupa huruf.

### Contoh

Isikan data tentang nama siswa, jenis kelamin dan tinggi badan pada table berikut kedalam data editor.

No.	Nama	Jenis Kelamin	Berat badan
1.	Muhammad mustain	Laki-laki	65 kg
2.	Faizatul abadiyah	Perempuan	59 kg
3.	Alisha almira ramadhani	Perempuan	25 kg
4.	Muhammad ahsanul haq	Laki-laki	35 kg
5.	Azwar Habibi	Laki-laki	80 kg

Penyelesaian :

Variable pertama adalah nama siswa, maka klik var1 kemudian kemudian dengan sendirinya secara otomatis variable editor isikan nama siswa pada variable nama dan pilih type data Character (karena nama siswa berupa huruf).

Variable kedua adalah jenis kelamin dengan type data character.

Variabel ketiga adalah berat badan dengan type data numeric (karena tinggi badan berupa angka).

Selanjutnya lakukan pengisian data sesuai dengan penamaan variabelnya, setelah semua data selesai diisikan, maka tutup tampilan data editor untuk mengakhiri proses input data. Data akan secara otomatis disimpan dengan nama latihan 1.

The screenshot shows the R Data Editor window titled "Data Editor: latihan1". It contains a table with the following data:

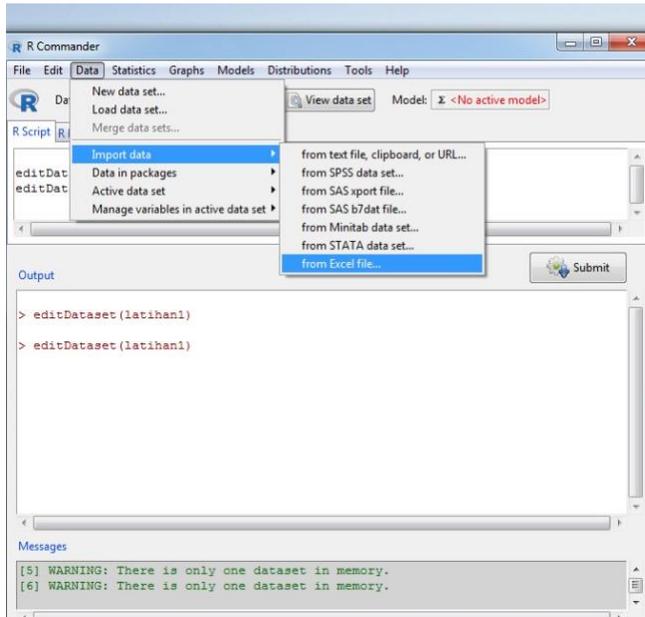
rowname	1	2	3	4
rowname	NAMA	JENIS KELAMIN	BERAT BADAN	V4
1	MUHAMMAD MUSTAIN	LAKI-LAKI	65	NA
2	FAIZATUL ABADIYAH	PEREMPUAN	59	NA
3	ALISHA ALMIRA RAMADHANI	PEREMPUAN	25	NA
4	MUHAMMAD AHSANUL HAQ	LAKI-LAKI	35	NA
5	AZWAR HABIBI	LAKI-LAKI	80	NA
6	NA	NA	NA	NA
7	NA	NA	NA	NA
8	NA	NA	NA	NA
9	NA	NA	NA	NA
10	NA	NA	NA	NA
11	NA	NA	NA	NA

#### 4. Impor data

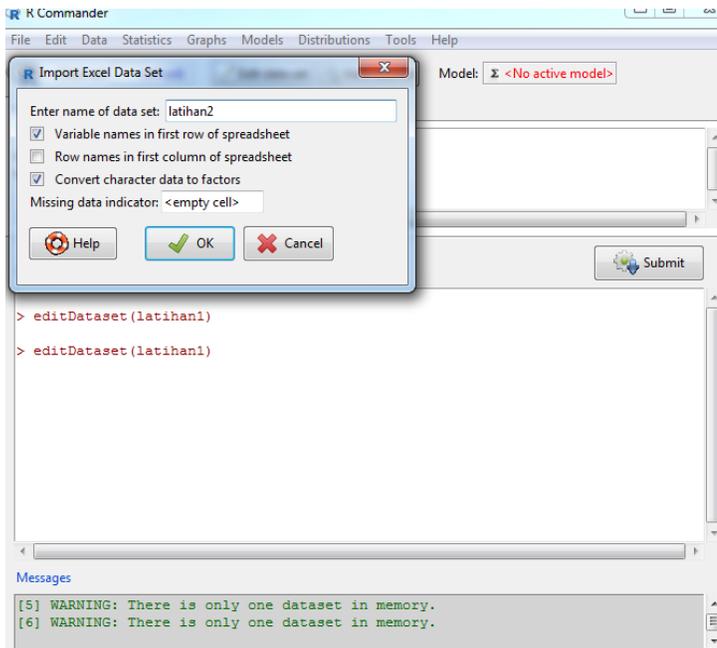
Didalam program R untuk paket R-commander juga difasilitasi dengan proses impor data dengan menggunakan menu klik impor data, yaitu dapat dilakukan dengan cara berikut:

Klik menu data -> impor data

Pilih program tempat penyimpanan file data yang akan diimpor, yaitu dari file teks atau clipboard, dataset SPSS, MINITAB, dataset STATA, dataset excel, Access atau dBase, seperti yang terlihat pada jendela impor data yaitu sebagai berikut:



Misalkan data diimpor dari data set excel, maka tampilan impor data set excel akan muncul. Masukkan nama untuk data set (misalkan latihan2) kemudian klik tombol Ok.



Kemudian arahkan ke folder tempat penyimpanan file excel yang akan diimpor, misalkan di folder D:/penulisan buku ajar/ dengan nama latihan2.xls kemudian klik open dan pilih sheet berapa data itu berada dalam ecell tersebut.

Data hasil impor dari file latihan2.xls akan disimpan ke dalam file latihan2.

a. Melihat data Set yang sedang aktif

Untuk melihat data yang sedang aktif saat ini, lakukan dengan mengklik tombol lihat data set. Kemudian, tampilan data editor akan dibuka dan menampilkan data set yang sedang aktif saat ini.

b. Memilih data set yang aktif

Untuk memilih data yang ingin diaktifkan pada R-commander dapat dilakukan dengan cara, mengklik menu data -> data set aktif -> pilih Data set Aktif...

Pilihlah data set yang ingin diaktifkan dengan mengklik nama data set yang dipilih, kemudian klik ok.

c. Mengedit data

Untuk mengedit data yang sudah kita inputkan dapat dilakukan dengan mengklik menu tombol edit data set, selanjutnya tampilan menu data editor akan dibuka dan proses edit data bias langsung dilakukan. Jika sudah selesai melakukan edit data, maka tutup tampilan menu data editor dan data yang sudah diedit akan tersimpan secara otomatis.

### C. Rangkuman

R adalah salah satu *software* analisis data yang termasuk kelompok *software* statistik yang *open source* yang tidak memerlukan lisensi atau gratis, yang dikenal dengan *freeware*. R dapat diperoleh secara gratis di <https://cran.r->

*project.org. software R* ini sudah Dilengkapi dengan banyak kemampuan internal untuk menganalisis data dan menampilkan grafik sehingga software R bisa dikategorikan sebagai paket pengolahan data (software statistika).

Beberapa kemampuan yang menonjol dari software R yang menjadi alasan banyak statistisi memilihnya sebagai program aplikasi antara lain yaitu, program R memiliki koleksi program analisis data yang disebut program R memiliki koleksi program analisis data, yang disebut library atau pustaka yang sangat luas seperti statistika deskriptif, regresi, pemodelan statistika (baik linier maupun non linier), Anova dan multivariat. variasi penampilan grafiknya nya sangat banyak dan berkualitas tinggi, baik penampilan di layar monitor maupun dalam bentuk cetak di atas. Kemampuan pemrograman (bahasa S) dapat dikembangkan secara fleksibel untuk kepentingan khusus yang lebih lanjut. Merupakan pemrograman yang berorientasi pada objek. keuntungannya, apabila apa yang telah dikerjakan R saat ini diperlukan di kemudian hari maka R dapat mengambilnya tanpa harus melakukan perhitungan ulang dari awal.

#### **D. Tugas**

Lakukan instalasi software R dan aktifkan paket R-Commander sesuai yang sudah dilakukan pada bab diatas, kemudian berikan laporan dari hasil penginstallan anda.

## **E. Rujukan**

- Kuswardi dan Mutiara, Erna. 2004. *Statistik Berbasis Komputer untuk Orang-orang Nonstatistik*. Jakarta : PT Alex Media Komputindo
- Maindonald, J.H. (2004). *Using R for Data Analysis and Graphics: Introduction, Code and Commentary*. Centre for Bioinformatics Science, Australian National University.
- Owen, W.J. (2007). *The R Guide*. Department of Mathematics and Computer Science, University of Richmond.
- Paradis, E. (2005). *R for Beginners*. Institut des Sciences de l'Evolution, Universite Montpellier II, France.
- Suhartono, (2008). *Analisis Data Statistik Dengan R*, Lab. Statistik Komputasi, ITS, Surabaya.
- Tirta, I. M. 2004. *Panduan R Pemrograman Untuk Analisis Data dan Grafik*. Jember: Laboratorium Statistika Jurusan Matematika FMIPA Universitas Jember.
- Venables, W.N. and Smith, D.M. (2007). *An Introduction to R*. The R Development Core Team.

## **F. Bacaan yang Dianjurkan**

- Kuswardi dan Mutiara, Erna. 2004. *Statistik Berbasis Komputer untuk Orang-orang Nonstatistik*. Jakarta : PT Alex Media Komputindo
- Maindonald, J.H. (2004). *Using R for Data Analysis and Graphics: Introduction, Code and Commentary*. Centre for Bioinformatics Science, Australian National University.
- Owen, W.J. (2007). *The R Guide*. Department of Mathematics and Computer Science, University of Richmond.
- Paradis, E. (2005). *R for Beginners*. Institut des Sciences de l'Evolution, Universite Montpellier II, France.
- Suhartono, (2008). *Analisis Data Statistik Dengan R*, Lab. Statistik Komputasi, ITS, Surabaya.

Tirta, I. M. (2004). *Panduan R Pemrograman Untuk Analisis Data dan Grafik*. Jember: Laboratorium Statistika Jurusan Matematika FMIPA Universitas Jember.

Venables, W.N. and Smith, D.M. (2007). *An Introduction to R*. The R Development Core Team.

# **BAB III**

## **STATISTIKA DESKRIPTIF**

### **A. Pendahuluan**

Pada Bab ini akan dibahas tentang ukuran pemusatan data, ukuran peyebaran data, ilustrasi grafik dan tabel, serta alokasinya dalam persoalan ekonomi dan bisnis.

#### **Tujuan Pembelajaran:**

Setelah mempelajari materi pada Bab ini mahasiswa diharapkan memahami ukuran pemusatan data, ukuran peyebaran data, ilustrasi grafik dan tabel, serta alokasinya dalam persoalan ekonomi dan bisnis.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nila  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Ukuran Pemusatan Data,
2. Ukuran Peyebaran Data,
3. Ilustrasi Grafik Dan Tabel,

### **B. Penyajian Data**

Statistik deskriptif merupakan bagian dari keilmuan statistik yang berkaitan dengan teknis untuk menjelaskan karakteristik data secara umum. Sebagai mana telah kita

ketahui, salah satu fungsi statistik adalah untuk mempermudah kita dalam mengolah dan menganalisis data hasil penelitian. Untuk dapat melakukan fungsi tersebut, maka data hasil penelitian yang telah diperoleh dari lapangan harus disajikan dalam bentuk yang sistematis dan informative. Sehingga dapat di lihat berapa nilai data yang paling kecil dan pling besar atau dapat diperoleh nilai rata-rata dari data tersebut.

Setiap melakukah kegiatan menganalisis data yang diperoleh dari penelitian, kita perlu menjelaskan beberapa karakteristik yang terdapat dari data tersebut. Dalam hal ini akan dilakukan dengan melihat ukuran deskriptif berupa angka atau nilai numeric yang sering digunakan untuk menjelaskan karakteristik data. Dikenal dengan dua jenis yaitu ukuran pemusatan data dan ukuran penyebaran data.

### 1. Ukuran Pemusatan Data

Ukuran pemusatan data atau yang sering dikenal dengan ukuran central tendensi merupakan suatu ukuran yang digunakan untuk melihat seberapa besar kecenderungan data terpusat pada nilai tertentu. Ukuran pemusatan data ini digunakan untuk menjelaskan lokasi pusat distribusi dari nilai-nilai pengamatan, pada bagian ini akan dibahas beberapa jenis ukuran pemusatan data yang sering digunakan mean. Mean atau rata-rata dari data merupakan nilai yang paling sering kita dapati dalam kehidupan sehari-hari. Rata-rata dari suatu set data didefinisikan sejumlah dari semua jumlah nilai pengamatan dibagi dengan jumlah data. Di notasikan

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

dengan rumus

Ukuran pemusatan data yang seanjutnya adalah median. Ukuran pemusatn ini erring digunakan untuk

menentukan nilai titik tengah setelah set data diurutkan. Sehingga median dari suatu set data didefinisikan nilai pengamatan yang terletak di tengah-tengah ketika data diurutkan dari yang terbesar ke yang paling kecil. Untuk suatu set data yang kecil, median terletak ditengah urutan, jika datanya ganjil. Sedangkan, untuk data yang jumlahnya genap median dihitung sebagai rata-rata dari dua data yang terletak di tengah-tengah urutan set data.

## 2. Ukuran Penyebaran Data

Semakin besar data yang dimiliki menunjukkan bahwa data yang bersangkutan semakin beragam. Keanekaragaman tersebut ditunjukkan dengan ukuran penyebaran data. Ukuran penyebaran data yang sering digunakan adalah standart deviasi atau simpangan baku. Secara matematis dapat dinyatakan sebagai akar jumlah deviasi kuadrat dibagi banyaknya individu dalam suatu distribusi. Simbul dari standart deviasi biasanya ditulis

$$SD = \sqrt{\frac{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}}{n - 1}}$$

SD sehingga dapat dirumuskan

Ukuran penyebaran data yang lain biasa disebut ukuran penempatan atau ukuran lokasi. Apabila pada ukuran pemusatan data untuk nilai tengah dikenal dengan median artinya pada nilai median ini membagi data menjadi 2 bagian yang sama, artinya 50% dari kelompok data menjadi nilai yang sama atau lebih kecil dari nilai-nilai median sedangkan 50% lainnya mempunyai nilai yang sama atau lebih besar dari nilai median. Sedangkan untuk ukuran penempatan yang lain didalam statistic yang paling sering digunakan adalah kuartil.

Kuartil merupakan nilai atau angka yang membagi data dalam empat bagian yang sama setelah data disusun atau diurutkan dari data terkecil sampai data yang terbesar atau sebaliknya. Ada tiga kuartil yaitu kuartil pertama, kuartil kedua dan kuartil ke tiga.

Kuartil pertama ialah nilai dalam distribusi yang membatasi 75% frekuensi dibagian atas dan 25% frekuensi dibagian bawah distribusi. Kuartil kedua yaitu nilai dalam distribusi yang membatasi 50% distribusi frekuensi di bagian atas dan 50% frekuensi dibagian bawah distribusi. Kuartil ketiga ialah nilai dalam distribusi yang membatasi 25% frekuensi dibagian atas dan 75% frekuensi di bagian bawah.

### **3. Ringkasan Data Menggunakan R-Commander**

Statistik deskriptif adalah metode statistika yang digunakan untuk mendeskripsikan atau menggunakan obyek penelitian melalui data sampel tanpa melakukan analisis dan kesimpulan. Dalam statistic deskriptif dikemukakan cara-cara penyajian data dengan grafik atau diagram baik itu diagram garis, diagram batang, dan diagram daun. Diagram lingkaran dan histogram dan lainnya. Selain itu juga disajikan ringkasan (summary) data seperti mean, median, modus serta keragaman kelompok seperti rantang (range atau jangkauan) dan simpangan baku (standart deviasi).

Pada bab ini akan dibahas tentang penggunaan R-Comannder untuk membuat statistic deskriptif dari suatu kumpulan data, khususnya pembuatan ringkasan (summary). Secara umum ada dua macam data yang akan dibuat ringkasan (summary), yaitu data numeric (metric) dan data nonmetric. Data metric berskala interval atau rasio, sedangkan data nonmetric berskala nominal atau ordinal.

Pada data metric (numeric), ringkasan yang akan ditampilkan adalah beberapa besaran statistic yaitu min, mean, kuartil 1, median, kuartil 3 dan max. sedangkan pada data nonnumerik ringkasan yang akan ditampilkan hanya jumlah atau frekuensi pada masing-masing kategori yang ada.

Pada contoh dibawah ini akan dilakukan statistic deskriptif dengan data yang disajikan yaitu tentang data hasil penjualan dan biaya produksi dari jenis usaha koperasi dan UMKM (Usaha Mikro Kecil Menengah).

Langkah pertama mengaktifkan data set yang akan digunakan untuk membuat ringkasan dari semua variable pada data set, yaitu variable numeric dan variable nonnumeric. Prosedur yang dilakukan menggunakan R-Commander yaitu dengan mengklik menu Statistika, pilih Summaries kemudian pilih active data set. Misalkan akan dibuat ringkasan dari semua variable pada data Ringkasan yang sebelumnya telah disimpan sebagai file Excel, maka langkah-langkah yang dilakukan untuk membuat ringkasan dari semua variable menggunakan R-Commander adalah :

- Bukalah file data ringkasan yang telah disimpan sebagai file Excel (lakukan impor data), yaitu :
  - ✓ Klik menu data → impor data → from excel file...
  - ✓ Masukkan nama untuk dataset (misalkan) latstatdes kemudian klik ok.
  - ✓ Selanjutnya arahkan ke direktori tempat penyimpanan file Excel yang akan diimpor kemudian klik open pilih sheet mana data itu berada di file excel tersebut.
  - ✓ Pada tampilan dialog R-Commander terlihat data set latstatdes sedang aktif

- ✓ Klik tombol view data set untuk melihat hasil impor data tersebut.

	jenis.usaha	biaya.produksi	penjualan
1	koperasi	550000	1250000
2	koperasi	600000	1350000
3	koperasi	500000	1200000
4	koperasi	450000	1100000
5	koperasi	550000	1405000
6	koperasi	555000	1250000
7	koperasi	505000	1350000
8	koperasi	600000	1300000
9	UMKM	855000	1750000
10	UMKM	755000	1850000
11	UMKM	800000	1750000
12	UMKM	700000	1700000
13	UMKM	850000	1750000
14	UMKM	900000	1850000
15	UMKM	800000	1800000
16	UMKM	750000	1750000
17	UMKM	850000	1800000
18	UMKM	900000	1900000
19	UMKM	800000	1855000
20	UMKM	750000	1850000
21	UMKM	700000	1800000
22	UMKM	855000	1900000
23	UMKM	950000	1855000
24	UMKM	900000	1850000
25	UMKM	955000	2000000

- ✓ Klik menu statistics → summaries → active data set  
Output yang diperoleh dari ringkasan semua variable adalah sebagai berikut :

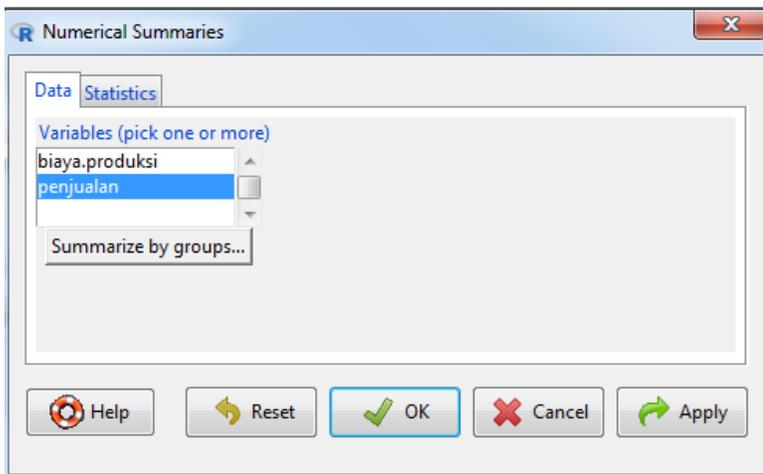
```
> summary(latstatdes)
jenis.usaha biaya.produksi      penjualan
koperasi: 8   Min.   :450000   Min.   :1100000
UMKM      :17  1st Qu.:600000   1st Qu.:1350000
           Median :755000   Median :1750000
           Mean   :735200   Mean   :1648600
           3rd Qu.:855000   3rd Qu.:1850000
           Max.   :955000   Max.   :2000000
```

Ringkasan numerik digunakan untuk membuat ringkasan dari suatu variabel numerik (metrik). Prosedur

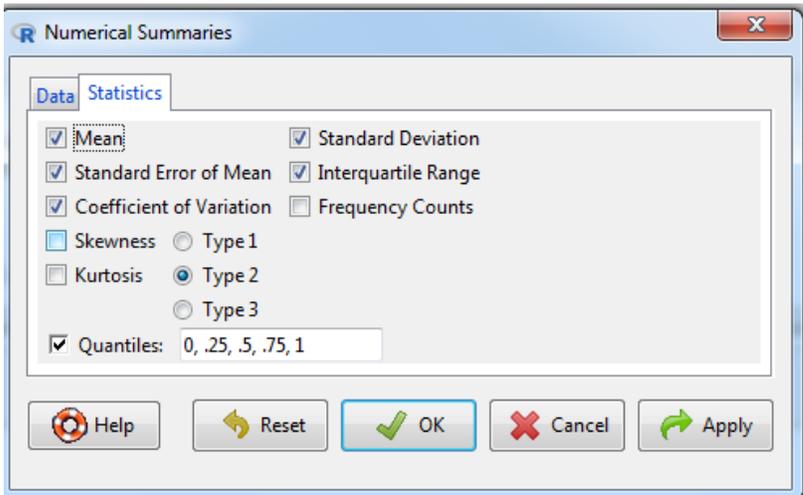
yang dilakukan menggunakan R-Commander yaitu dengan mengklik menu statistics, pilih summaries dan kemusial pilih numerical summaries...

Misalkan akan dibuat ringkasan numerik dari data ringkasan yang sudah disimpan dalam file excel yaitu latstatdes. Langkah-langkah yang dilakukan untuk membuat numerical summaries menggunakan R-Commander adalah :

- Klik menu statistics→summaries→numerical summaries...
- Tampilan pilihan variable numeric yang akan dibuat ringkasan numeriknya akan muncul seperti berikut :



- pilih salah satu variable numeric yang akan dibuat ringkasan numeriknya pada pilihan variable. Misalkan pilih variable penjualan.
- Pilih atau centang besaran-besaran atatistik yang akan ditampilkan dalam hasil ringkasannya. Misalkan akan ditampilkan nilai Mean, standart error of mean, standart deviasi dan lain-lain.
- Klik tombol OK.



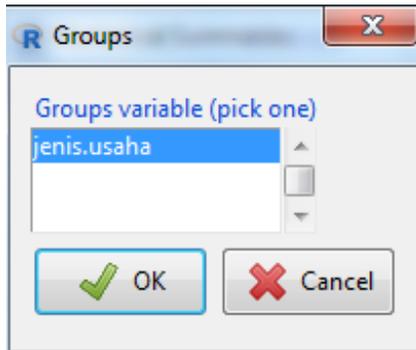
Output yang diperoleh dari ringkasan numeric adalah sebagai berikut :

```
> numSummary(latstatdes[, "penjualan", drop=FALSE],
  statistics=c("mean", "sd", "se(mean)", "IQR", "quantiles",
"cv"), quantiles=c(0, .25, .5, .75, 1))
      mean      sd se(mean)   IQR      cv      0%      25%
50%      75%      100%  n
1648600 273025.2 54605.04 500000 0.1656103 1100000 1350000
1750000 1850000 2000000 25
```

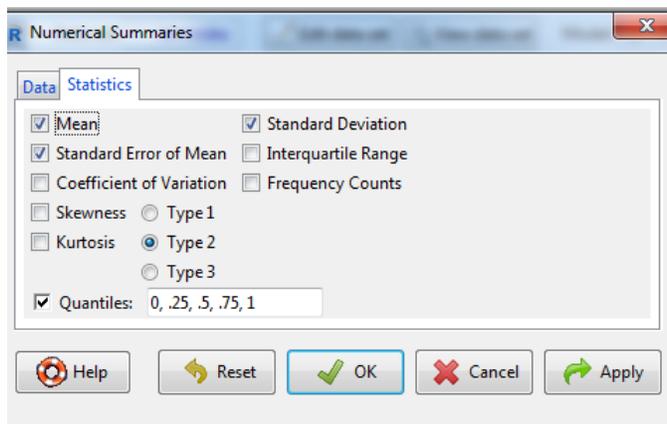
Apabila ingin membuat ringkasan numeric berdasarkan kelompok atau factor tertentu, maka pada tampilan menu pilihan variable numeric yang muncul seperti pada gambar dibawah ini. Misalkan akan dibuat ringkasan numeric untuk variable jenis usaha. Maka langkah langkah yang akan dilakukan untuk membuat ringkasan numeric berdasarkan kelompok atau factor tertentu menggunakan R-Commander adalah :

- Klik menu statistic → summaries → numerical summaries...
- Pada tampilan menu pilihan variable numeric yang muncul, pilih variable Janis usaha pada pilihan variable kemudian klik summarize by groups...
- Pilih variable jenis usaha pada pilihan groups variable..

- Klik OK



- Pilih atau centang besaran-besaran atastistik yang akan ditampilkan dalam hasil ringkasannya. Misalkan akan ditampilkan nilai Mean, standart error of mean, standart deviasi dan lain-lain.
- Klik ok



Output yang diperoleh dari ringkasan numeric untuk variable nenis usaha adalah sebagai berikut :

```
> numSummary(latstatdes[, "penjualan", drop=FALSE],
groups=latstatdes$jenis.usaha, statistics=c("mean", "sd",
"se(mean)", "quantiles"), quantiles=c(0, .25, .5, .75, 1))
      mean      sd se(mean)      0%      25%      50%
75%    100% penjualan:n
koperasi 1275625 97300.62 34400.96 1100000 1237500 1275000
1350000 1405000           8
UMKM      1824118 73383.31 17798.07 1700000 1750000 1850000
1855000 2000000           17
```

#### 4. Ilustrasi Grafik Dan Tabel

Penyajian data yang terakhir yaitu secara grafik dan tabel. Disamping sudah dijelaskan diatas Suatu bentuk tulisan/narasi dalam penjelasan data adapula secara tabel dan gambar/grafik yang merupakan hasil kompilasi dan pengolahan data dengan Tujuan memberikan informasi dan memudahkan interpretasi hasil analisis sehingga Bentuk penyajian data yaitu secara TULISAN atau NARASI, TABEL atau DAFTAR dan secara GAMBAR atau GRAFIK atau DIAGRAM.

Penjelasan secara TULISAN atau NARASI digunakan untuk membuat laporan. Contohnya Luas wilayah kerja Puskesmas A 40 km<sup>2</sup>, Penduduk di wilayah kerja Puskesmas A 54.000 jiwa, Daerah kerja Puskesmas A terdiri dari 10 (sepuluh) desa, dengan rincian yaitu 2 Puskesmas Pembantu (Pustu), 6 Polindes dan 1 Puskesmas Keliling.

Penyajian secara tabel atau daftar yaitu Penyajian dalam bentuk angka (data numerik) yang disusun dalam kolom dan baris. Dengan Tujuannya menunjukkan frekuensi kejadian dalam kategori yang berbeda. Jenis tabel Tabulasi satu variabel (*One-Way Frequency Tables*) atau tabulasi distribusi frekuensi, Tabulasi silang (*Two-Way Frequency Tables*) atau *text table*, disebut juga tabel kontingensi dan Tabulasi stratifikasi (*Three-Way Frequency Tables*).

Syarat Pembuatan Tabel

- Judul tabel : memberikan keterangan yang mandiri tentang "apa/siapa", dimana dan kapan dilakukan
- Terdiri dari baris dan kolom
- Ada badan data yang menerangkan tentang frekuensi distribusi kejadian

- Ada jumlah kolom dan jumlah baris serta jumlah keseluruhan (total silang dari jumlah kolom dan jumlah baris)
- Ada catatan kaki (*footnote*) untuk menerangkan sumber data diperoleh atau keterangan lain yang diperlukan untuk memperjelas penyajian tabel, misalnya kode atau singkatan Bentuk umum tabel dan Judul Tabel.

Judul Baris	Judul Kolom		Jumlah Baris
	sel	sel	
Jumlah Kolom			Jumlah total

- Pengklasifikasian terhadap data kuantitatif, perhatikan :
  - a. Jumlah kelas (5-15) :
    - jumlah pengamatan
    - tujuan
    - perjanjian
  - b. Pengamatan terkecil dan terbesar harus masuk, tidak ada *overlapping*
  - c. Lebar kelas interval harus sama

Rumus Sturge : 
$$i = \frac{r}{1 + 3,3 \log n}$$

Rumus lain : 
$$i = \frac{r}{c}$$

- d. Hindari (bila mungkin) *Open-ended Class*
  - e. Usahakan pemusatan data jatuh pada *Midpoint Class Interval*
1. Tabel distribusi frekuensi

- bentuk tabel sederhana, biasanya data terdiri dari 1 variabel disertai frekuensi masing-masing kategori variabel tersebut
- terdiri dari 2 kolom :
  - \* kolom pertama berisi kategori atau kelompok dari variabel yang diamati dan dicatat
  - \* kolom kedua berisi frekuensi dari masing-masing kategori

Kelompok umur	Frekuensi	%
0 - 4	35	8,75
5 - 14	45	11,25
15 - 24	125	31,25
25 - 34	95	23,75
35 - 44	100	25,00
Jumlah	400	100,00

#### Tabel kontingensi (tabulasi silang)

- Untuk data yang terdiri dari dua variabel atau dua faktor, dibuat tabel kontingensi atau tabulasi silang berukuran  $b \times k$  dengan  $b$  menyatakan banyak baris dan  $k$  menyatakan banyak kolom
- Untuk mengetahui hubungan antar variabel atau membedakan proporsi suatu kejadian antar kelompok yang berbeda
- Secara umum, variabel independen diletakkan pada kolom dan variabel dependen pada baris
- Pemberian persentase menurut masing-masing kolom

#### Contoh Tabulasi Silang

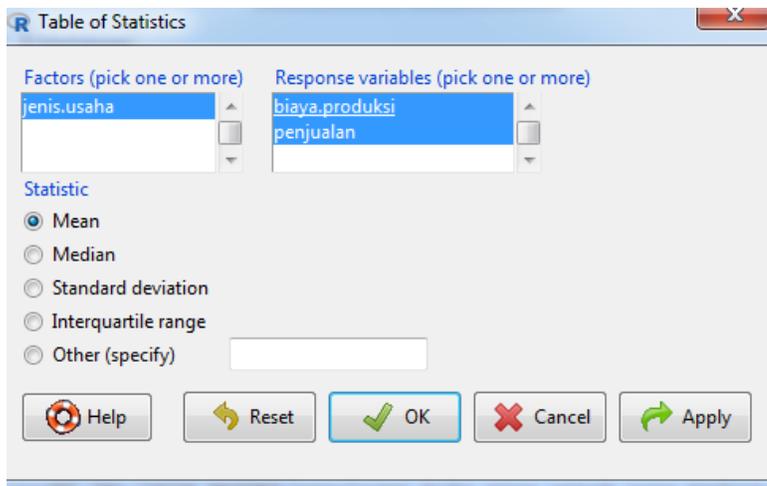
Triwulan	Rata-rata cakupan indeks pertumbuhan ekonomi (dalam %)	
	1996	1997
I	56,1	45,0
II	73,2	80,0
III	23,4	44,2
IV	33,5	35,4

Tabel statistika digunakan untuk membuat ringkasan statistic dalam bentuk tabel untuk suatu variabel numeric (*metric*) berdasarkan variable nonnumerik (*nonmetric*) atau faktor tertentu.

Prosedur yang dilakukan menggunakan R-Commander yaitu dengan mengklik menu, pilih summaries dan kemudian pilih *table of statistics...*

Misalkan akan dibuat tabel statistika untuk variabel jenis usaha berdasarkan variabel penjualan dan Biaya produksi. Dari data yang ada sudah disimpan dalam file Excel yaitu *latstatdes*. langkah-langkah yang dilakukan untuk membuat tabel statistika menggunakan R-Commander adalah

- Klik menu *statistic* → *summaries* → *table of statistics...*
- Tampilan pilihan factor dan variable respon yang akan ditampilkan tabel statistiknya akan muncul, seperti pada gambar berikut :



- Pilih variable jenis usaha pada variable factor dan pilih variable biaya produksi dan penjualan pada variable respon.

- Pilih besaran statistic yang akan ditampilkan pada tabel statistika yang akan dibuat. Misalkan akan ditampilkan mean.
- Klik tombol ok.

Output yang diperoleh dari tabel statistika untuk rata-rata jenis usaha berdasarkan penjualan dan biaya produksi adalah sebagai berikut :

```
> Tapply(biaya.produksi ~ jenis.usaha, mean,
na.action=na.omit,
+ data=latstatdes) # mean by groups
koperasi      UMKM
538750.0 827647.1
```

```
> # Table for penjualan:
```

```
> Tapply(penjualan ~ jenis.usaha, mean,
na.action=na.omit,
+ data=latstatdes) # mean by groups
koperasi      UMKM
1275625 1824118
```

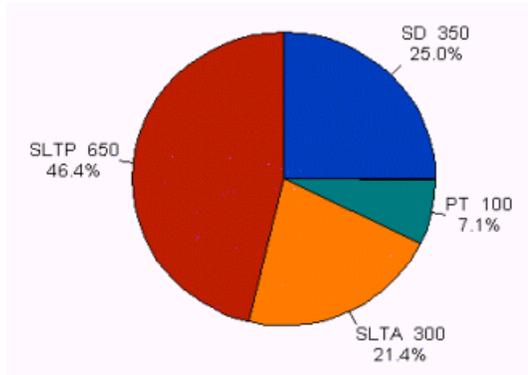
R menyediakan banyak menu pilihan grafik untuk penyajian data dalam statistik deskriptif melalui R-Commander , antara lain yaitu histogram , diagram batang dan daun, diagram batang diagram lingkaran dan lain-lain.

#### i. DIAGRAM LINGKARAN (PIE DIAGRAM)

menyajikan data kualitatif sebagai bagian komponen perbandingan dari keseluruhan

Syarat :

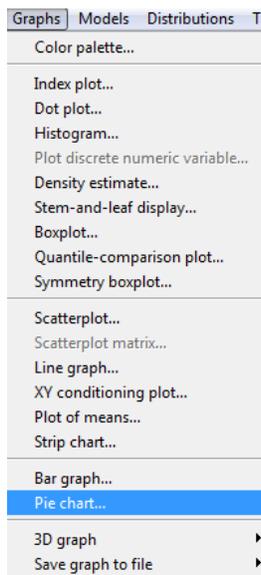
- Bentuk lingkaran dengan jumlah komponen 100 % atau 360°
- Perhitungan luas komponen atau sektor merupakan perbandingan yang dikalikan dengan 100%



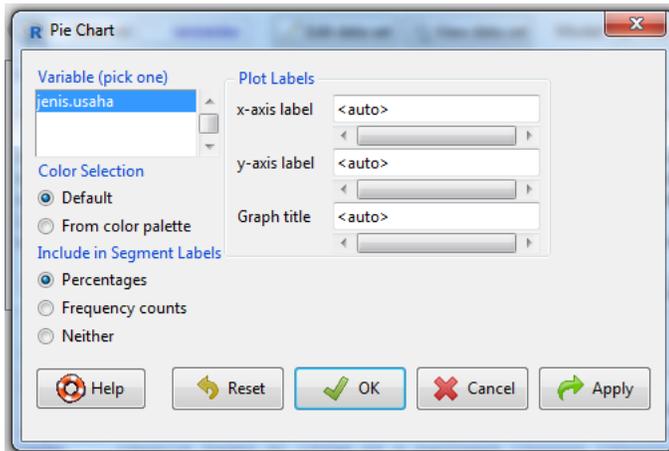
Dalam program R untuk menu diagram Lingkaran yaitu menu graph kemudian pilih pie chart...

Misalkan akan dibuat diagram lingkaran untuk variable jenis usaha dari data yang ada dan sudah disimpan dalam file R, yaitu latstatdes, prosedur pembuatan diagram lingkaran menggunakan Rcommander adalah :

- Klik menu graph → pie chart...

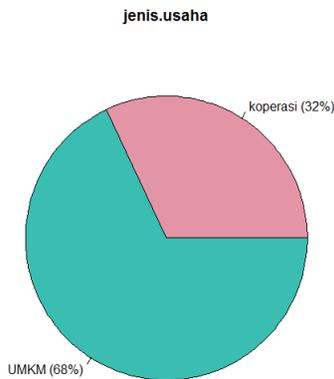


- Tampilan pilihan variable dalam pembuatan diagram lingkaran akan muncul, seperti berikut



- Pilih variable jenis usaha pada pilihan variable
- Kemudian klik ok

Output diagram lingkaran yang diperoleh adalah sebagai berikut :



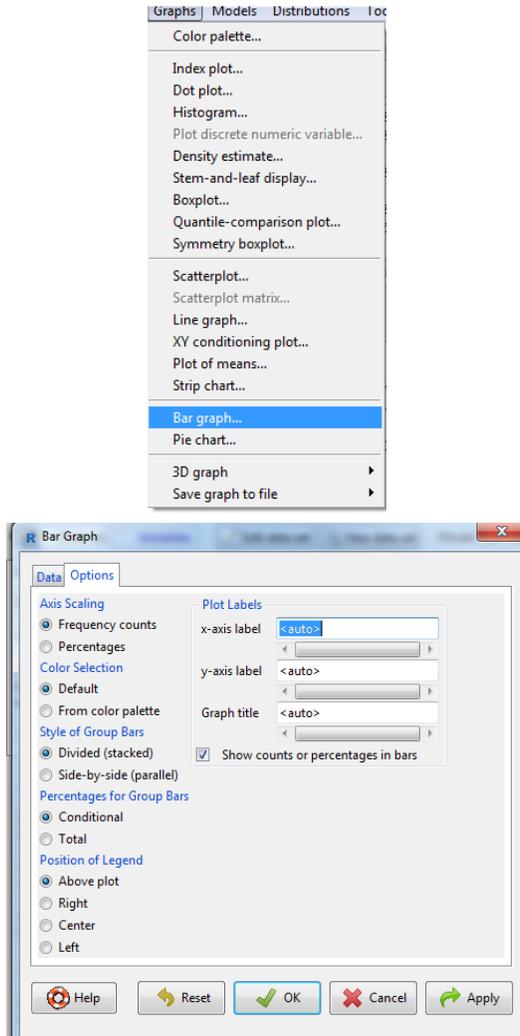
## ii. DIAGRAM BATANG

Bertujuan melihat kecenderungan data pengamatan menurut waktu, dimana sumbu x berisi data waktu dan sumbu y menunjukkan frekuensi nilai dari variabel data.

Menu yang digunakan untuk membuat diagram batang melalui R-Commander adalah menu Graphskemudian pilih Bar Graph...

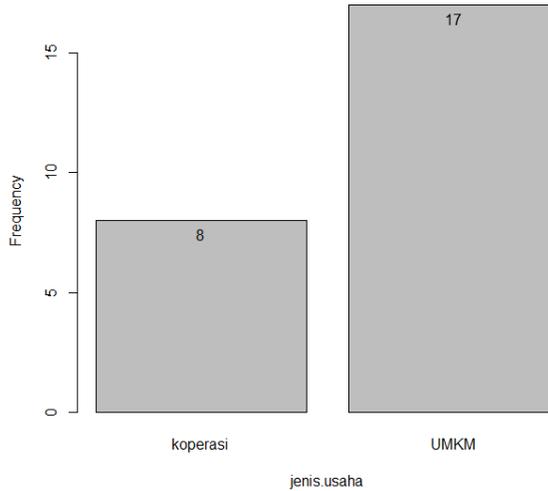
Misalkan akan dibuat diagram batang untuk variabel jenis usaha dari data yang ada dan sudah tersimpan di Program R yaitu latstasdes. Prosedur pembuatan diagram batang menggunakan R-Commander adalah :

- Klik menu graphs→bar graph...
- Tampilan pilihan variabel dalam pembuatan histogram akan muncul, seperti berikut



- Pilih variable jenis usaha pada pilihan variable.
- Kemudian klik OK.

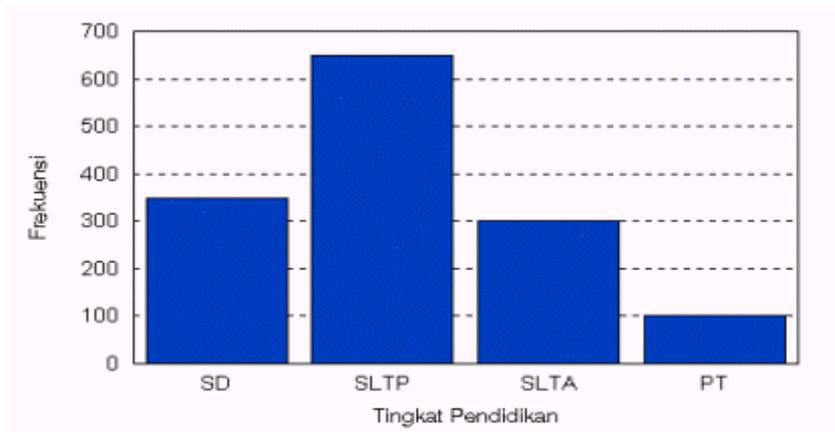
Output diagram batang yang diperoleh adalah :



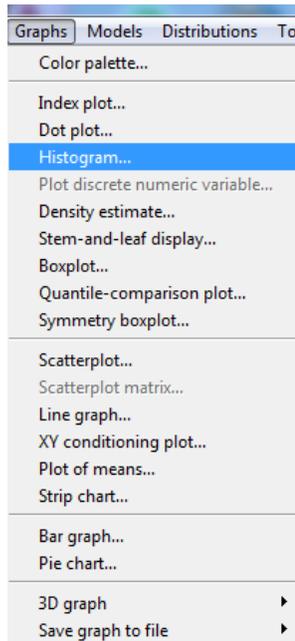
membandingkan beberapa pengamatan data menurut tempat dan jenis atau kategori tertentu

Macam :

1. Histogram

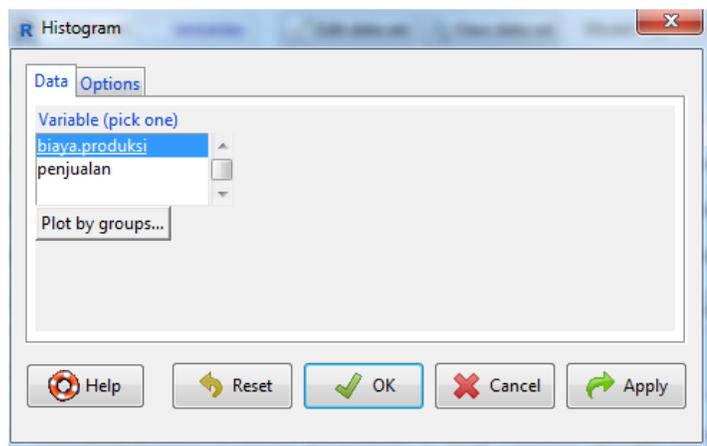


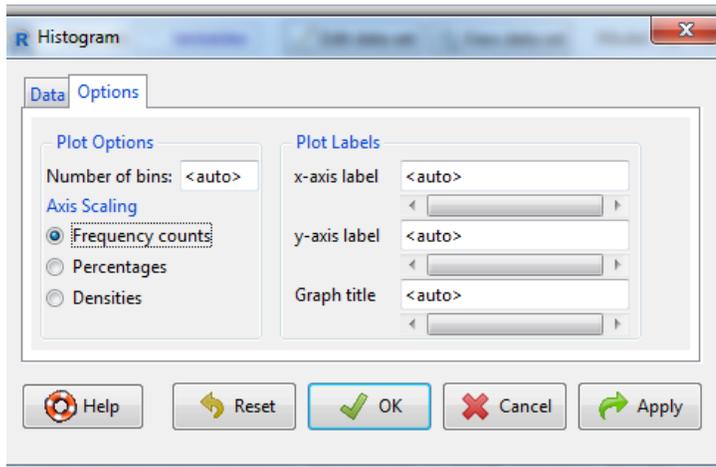
Menu Yang digunakan untuk membuat histogram melalui R-Commander adalah menu Graph kemudian pilih Histogram....



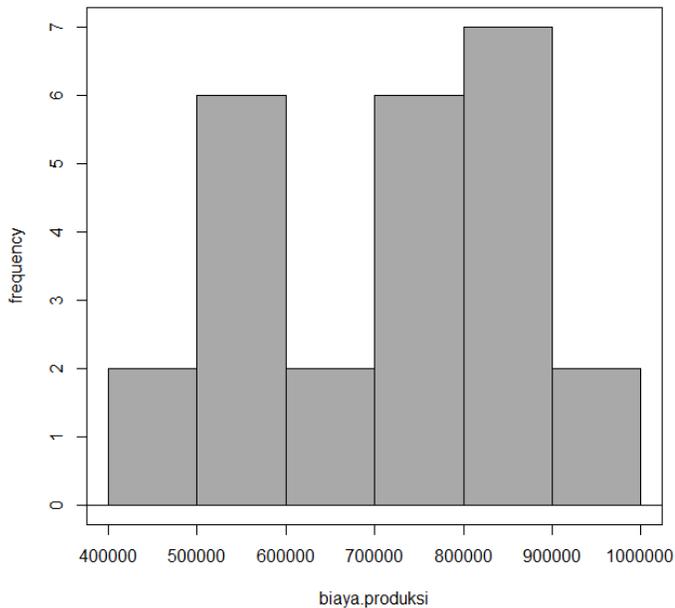
Misalkan akan dibuat histogram untuk variable penjualan dari data ringkasan yang sudah disimpan dalam R, yaitu `latstatdes`. Prosedur pembuatan histogram menggunakan R-Commander adalah :

- Klik menu `graphs` → `histogram...`
- Tampilan pilihan variabel dalam pembuatan histogram akan muncul, seperti berikut





- Pilih variable biaya produksi pada pilihan variable yang akan dibuat histogramnya.
- Pilih skala sumbu yang diinginkan. Misalkan pilih frequency
- Kemudian klik OK.  
Output histogram yang diperoleh adalah :



### C. Rangkuman

Statistik deskriptif merupakan bagian dari keilmuan statistik yang berkaitan dengan teknis untuk menjelaskan karakteristik data secara umum. dua jenis yaitu ukuran pemusatan data dan ukuran penyebaran data. Ukuran pemusatan data atau yang sering dikenal dengan ukuran central tendensi merupakan suatu ukuran yang digunakan untuk melihat seberapa besar kecenderungan data terpusat pada nilai tertentu.

Rata-rata dari suatu set data didefinisikan sejumlah dari semua jumlah nilai pengamatan dibagi dengan jumlah

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

data. Di notasikan dengan rumus  $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ . Ukuran pemusatan data yang selanjutnya adalah median nilai tengah setelah data diurutkan. Semakin besar data yang dimiliki menunjukkan bahwa data yang bersangkutan semakin beragam. Keanekaragaman tersebut ditunjukkan dengan ukuran penyebaran data. Ukuran penyebaran data yang sering digunakan adalah standart devisu atau simpangan baku. Simbul dari standart deviasi biasanya ditulis SD sehingga

$$SD = \sqrt{\frac{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}}{n-1}}$$

dapat dirumuskan  $SD = \sqrt{\frac{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}}{n-1}}$ . Ukuran penyebaran data yang lain biasa disebut ukuran penempatan atau ukuran lokasi.

Penggunaan perhitungan dengan bantuan program R disajikan ringkasan (summary) data seperti mean, median, modus serta keragaman kelompok seperti rantang (range atau jangkauan) dan simpangan baku (standart deviasi).

Penyajian data yang terakhir yaitu secara grafik dan tabel. Disamping sudah dijelaskan diatas Suatu bentuk tulisan/narasi dalam penjelasan data adapula secara tabel dan gambar/grafik yang merupakan hasil kompilasi dan pengolahan data dengan Tujuan memberikan informasi dan memudahkan interpretasi hasil analisis sehingga Bentuk penyajian data yaitu secara TULISAN atau NARASI, TABEL atau DAFTAR dan secara GAMBAR atau GRAFIK atau DIAGRAM.

#### **D. Tugas**

Lakukan analisis statistik deskriptif dari hasil pencarian data yang sudah dilakukan, kemudian berikan analisis dari hasil interpretasi anda.

#### **E. Rujukan**

- Aridinanti, L dkk. (2003). *Pengantar Metoda Statistika*. Surabaya : Jurusan Statistika FMIPA ITS.
- Purwanto. (2003). *Ekonomi Untuk Keuangan dan Bisnis*. Jakarta: Salemba Empat
- Walpole, E R. (1995). *Pengantar Statistik Edisi Ketiga*. Jakarta : PT. Gramedia Pustaka Utama

#### **F. Bacaan yang Dianjurkan**

- Aridinanti, L dkk. (2003). *Pengantar Metoda Statistika*. Surabaya : Jurusan Statistika FMIPA ITS.
- Purwanto. (2003). *Ekonomi Untuk Keuangan dan Bisnis*. Jakarta: Salemba Empat
- Walpole, E R. (1995). *Pengantar Statistik Edisi Ketiga*. Jakarta : PT. Gramedia Pustaka Utama

# BAB IV

## PENGUJIAN HIPOTESIS

### A. Pendahuluan

Pada Bab ini akan dibahas konsep uji hipotesis dan bentuk hipotesis serta aplikasinya dalam persoalan ekonomi dan bisnis.

#### **Tujuan Pembelajaran :**

Setelah mempelajari materi pada Bab ini mahasiswa diharapkan memahami konsep uji hipotesis dan bentuk hipotesis serta aplikasinya dalam persoalan ekonomi dan bisnis.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nilai  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. konsep uji hipotesis
2. bentuk hipotesis

### B. Konsep Uji Hipotesis

Perkembangan ilmu sosial ditunjang oleh penelitian (*research*) dalam menentukan kesimpulan dari penelitian yaitu dengan melakukan penerimaan hipotesis dari teori yang diturunkan dari teori yang diturunkan dari teori-teori

sosialnya. Setelah memilih hipotesis dari teori yang memungkinkan, selanjutnya adalah mengumpulkan data yang dapat memberikan keterangan langsung dalam penentuan hipotesis tersebut. Keputusan dari data yang memberikanketerangan atau kesimpulan akan memberikan pengukuhan, perbaikan atau penolakan terhadap teori dari sumbernya.

Dengan tujuan untuk mendapatkan keputusan yang objektif, seperti misalnya apakah hipotesis tersebut ditunjang oleh data yang cukup, sehingga bias menentukan apakah hipotesis tersebut ditolak atau diterima, obyektif sangat diperlukan sebab misalnya kesimpulan yang diperoleh dari penelitian akan sampai kemasyarakat dan juga ada kemungkinan penelitian ini akan diteruskan oleh peneliti yang lain. Langkah-langkah yang objektif harus didasari oleh keterangan-keterangan yang didapat peneliti. Jika tidak demikian resikonya, kesimpulan atau keputusan terhadap hipotesis akan tidak terkontrol.

Suatu hipotesis statistik adalah suatu pernyataan atau asumsi yang mungkin salah dan mungkin benar tentang parameter populasi. Kebenaran atau ketidak benaran suatu hipotesis tidak pernah diketahui secara pasti, kecuali jika kita memeriksa atau meneliti seluruh populasinya. Hal ini seringkali tidak mungkin dilakukan karena berbagai kendala, baik waktu, biaya, maupun tenaga yang harus dialokasikan untuk hal tersebut. Oleh karena itu, pemeriksaan terhadap kebenaran atau ketidak benaran hipotesis tersebut umumnya dilakukan melalui pengambilan sampel dari populasi.

Tujuan pengujian hipotesis adalah untuk memilih salah satu dari dua hipotesis tentang parameter populasi, yang keduanya saling bertentangan, yaitu hipotesis nol dinyatakan dengan  $H_0$ , dan hipotesis alternative atau hipotesis penelitian,

dinyatakan dengan  $H_1$ . Kedua hipotesis tersebut bersifat saling asing (mutually exclusive), artinya jika satu hipotesis ditolak, maka sebagai konsekuensinya, hipotesis lain diterima.

Pengujian hipotesis berdasarkan pada konsep pembuktian melalui pengkontradikian, yaitu jika analisis terhadap sampel menunjukkan ketidak konsistenan dengan hipotesis yang diuji, maka hipotesis tersebut ditolak dan disimpulkan bahwa hipotesis tersebut salah. Sebaliknya, jika analisis sampel tersebut konsisten dengan hipotesis yang diuji, maka hipotesis tersebut diterima. Namun demikian, penerimaan terhadap suatu hipotesis semata-mata sebagai akibat tidak cukupnya bukti untuk menolak hipotesis dan tidak berarti bahwa hipotesis tersebut adalah benar.

Prosedur pengujian hipotesis terdiri dari komponen komponen berikut ini:

- a. Hipotesis nol
- b. Hipotesis alternative atau hipotesis penelitian
- c. Statistik uji (metode analisis data)
- d. Daerah kritis atau daerah penolakan hipotesis

Hipotesis nol adalah asumsi atau anggapan yang berkaitan dengan nilai parameter populasi yang akan diuji. Hipotesis nol umumnya menyatakan bahwa nilai parameter populasi tersebut sama dengan suatu nilai tertentu. Misalkan  $A$  adalah parameter populasi, maka  $H_0$  umumnya dinyatakan dengan bentuk.

$H_0 : A = A_0$  dimana  $A_0$  adalah suatu nilai tertentu

Hipotesis alternative merupakan suatu pernyataan alternative jika asumsi atau anggapan tentang parameter populasi tersebut ternyata salah atau ditolak. Hipotesis alternative dapat mengambil salah satu dari ketiga bentuk berikut :

$H_0 : A \neq A_0$  atau  $H_0 : A < A_0$  atau  $H_0 : A > A_0$

Secara umum, terdapat tiga bentuk format pasangan hipotesis nol dan hipotesis alternative dalam pengujian hipotesis. Misalkan  $B$  adalah parameter populasi dan  $B_0$  adalah anggapan tentang nilai parameter tersebut, maka ketiga pasangan hipotesis tersebut adalah sebagai berikut :

- (i)  $H_0 : B = B_0$   
 $H_1 : B \neq B_0$  (hipotesis dua arah)
- (ii)  $H_0 : B = B_0$   
 $H_1 : B > B_0$  (hipotesis satu arah)
- (iii)  $H_0 : B = B_0$   
 $H_1 : B < B_0$  (hipotesis satu arah)

Untuk memilih salah satu dari kedua hipotesis tersebut ( $H_0$  atau  $H_1$ ) diperlukan suatu kriteria pengujian yang ditentukan berdasarkan pada suatu metode statistic uji. Penentuan statistik uji tersebut didasarkan atas statistic sampel dan distribusi smplingnya artinya berdasarkan bnyak atau sedikitnya sampel. Dengan demikian, satistik uji merupakan suatu varibel random yang nilai-nilainya digunakn untuk mengambil keputusan apakah menolak atau menerima hipotesis nol. Nilai-nilai statistic yang digunkan untuk menolak hipotesis nol disebut dengan ebagai daerah kritis atau daerah penolakan hipotesis. Sedangkan, nilai-nilai yang digunakan untuk menerima hipotesis nol disebut sebagai daerah penerimaan. Daerah penolakan dan daerah penerimaan hipotesis dibatasi oleh suatu nilai yang disebut dengan titik kritis.

Dalam setiap pengujian hipotesis, kita harus selalu memutuskan apakah menerima atau menolak hipotesis nol dan selalu ada kemungkinan bahwa kita membuat kesalahan dalam pengambilan keputusan tersebut. Keslahan tersebut terjadi ketika kita menolak suati hiotesis yang benar tau menerima hipotesis yang salah. Kedua jenis kesalahan ini diberinama secara khusus dalam pengujian hiotesis, yaitu

Kesalahan jenis I yaitu kesalahan ini terjadi ketika kita menolak  $H_0$  padahal  $H_0$  benar. Peluang terjadinya kesalahan ini dinyatakan dengan  $\alpha$ , dan disebut dengan taraf nyata (level of signficn)

Kesalahn tipe II yaitu kesalahan ini terjadi ketika kita menerima  $H_0$  padahal  $H_0$  salah dan  $H_1$  benar. Peluang terjadinya kesalahan ini dinyatakan dengan  $\beta$ . Sehingga komplemen dari  $\beta$ , yaitu  $(1 - \beta)$  disebut dengan kekuatan pengujian (*power of statistical test*).

Idealnya daerah penerimaan dan penolakan ditentukan agar meminimumkan  $\alpha$  dan  $\beta$  sekaligus., tetapi hal ini tidak mungkin dilakukan, karena peluang terjadinya kedua kesalahan tersebut berkaitan satu sama lainnya. Untuk suatu ukuran sampel tertentu., jika daerah penolakan dirubah agar memperkecil  $\alpha$ , maka  $\beta$  secara otomatis akan bertambah besar. Demikian sebaliknya. Jika  $\alpha$  diperbesar maka  $\beta$  akan mengecil.

Prosedur umum dilakukan oleh peneliti adalah dengan menentukan taraf nyata  $\alpha$  pada suatu nilai tertentu, hal ini akan secara otomatis menentukan nilai  $\beta$  bagi percobaan tersebut. Taraf nyata yang biasanya digunakan adalah  $\alpha$  sama dengan 0,1; 0,05 atau 0,01. untuk nilai  $\alpha$  tertentu, nilai  $\beta$  dapat diperkecil dengan memperbesar ukuran sampelnya.

Secara umum, prosedur Pengujian Hipotesis dilakukan melalui langkah-langkah berikut:

- (i) Tentukan  $H_0$  dan  $H_1$
- (ii) Dengan mengasumsikan bahwa  $H_0$  benar, Tentukan statistik uji berdasarkan atas distribusi sampling
- (iii) Tentukan daerah penolakan dan penerimaan  $H_0$  berdasarkan atas taraf nyata  $\alpha$ , bentuk  $H_1$  dan statistik uji nya
- (iv) Hitung nilai statistik uji dari sampel

(v) Ambil keputusan untuk menerima atau menolak  $H_0$  berdasarkan atas Hasil pengujian statistik uji

Langkah dalam menentukan statistik uji, merupakan salah satu langkah yang paling penting, karena Statistik tersebut digunakan untuk menentukan daerah penerimaan dan penolakan hipotesis, yang pada akhirnya digunakan untuk mengambil keputusan Apakah menerima atau menolak  $H_0$ . taraf nyata  $\alpha$  digunakan untuk menentukan titik kritis dari statistik uji tersebut, Sedangkan bentuk 1 digunakan untuk menentukan arah daerah kritis pengujian. jika merupakan suatu hipotesis dua arah, maka daerah kritis pengujian terletak di ujung-ujung kurva distribusi sampling statistik tersebut. jika  $H_1$  merupakan hipotesis satu arah, maka daerah kritis pengujian terletak di salah satu ujung kurva distribusi sampling yang bersesuaian dengan arah dari  $H_1$ . artinya, jika  $H_1$  mempunyai bentuk hipotesis lebih kecil dari ( $<$ ), maka daerah kritis pengujian terletak di ujung kiri kurva distribusinya, sedangkan jika  $H_1$  mempunyai bentuk hipotesis lebih besar dari ( $>$ ), maka daerah kritis pengujian terletak di ujung kanan kurva distribusinya.

#### 1. Bentuk Hipotesis

Hipotesis merupakan jawaban sementara terhadap rumusan masalah penelitian. dikatakan sementara karena, jawaban yang diberikan baru didasarkan pada teori dan belum menggunakan fakta.

Dalam penelitian yang menggunakan analisis statistik inferensial, terdapat dua hipotesis yang perlu diuji, yaitu hipotesis penelitian dan hipotesis statistik. menguji hipotesis penelitian berarti menguji jawaban yang sementara (tentatif) itu apakah betul-betul terjadi pada sampel yang diteliti atau tidak. kalau terjadi berarti hipotesis penelitian terbukti. selanjutnya menguji hipotesis statistik, berarti menguji Apakah hipotesis

penelitian yang telah terbukti atau tidak terbukti berdasarkan data sampel itu dapat diberlakukan pada populasi atau tidak.

Menurut tingkat penjelasan (level of explanation) variabel yang diteliti, Maka terdapat tiga bentuk hipotesis yang dirumuskan dan diuji, yaitu:

a. Hipotesis deskriptif

Hipotesis deskriptif, merupakan dugaan terhadap nilai satu variabel dalam satu sampel walaupun di dalamnya bisa terdapat beberapa kategori.

contoh

$H_0$  : kecenderungan masyarakat memilih warna mobil gelap

$H_1$  : kecenderungan masyarakat memilih warna mobil bukan warna gelap

b. hipotesis komparatif

Merupakan dugaan terhadap perbandingan nilai 2 sampel atau lebih. dan komparasi ini terdapat beberapa macam yaitu:

i) komparasi berpasangan related dalam dua sampel dan lebih dari 2 sampel k sampel

ii) komparasi independen dalam 2 sampel dan lebih dari 2 sampel k sampel

contoh:

sampel berpasangan, komparasi dua sampel

$H_0$  : Tidak terdapat perbedaan nilai penjualan sebelum dan sesudah ada iklan

$H_1$  : Terdapat perbedaan nilai penjualan sebelum dan sesudah ada iklan

sampel independen, komparasi tiga sampel

$H_0$  : Tidak terdapat perbedaan antara birokrat, akademisi dan pebisnis dalam memilih partai

$H_1$  : Terdapat perbedaan antara birokrat, akademisi dan pebisnis dalam memilih partai

iii) Hipotesis asosiatif (hubungan)

merupakan dugaan terhadap hubungan antara dua variabel atau lebih  
contoh

$H_0$  : tidak terdapat hubungan antara jenis profesi dan jenis olahraga yang disenangi

$H_1$  : terdapat hubungan antara jenis profesi dan jenis olahraga yang disenangi

### C. Rangkuman

Suatu hipotesis statistik adalah suatu pernyataan atau asumsi yang mungkin salah dan mungkin benar tentang parameter populasi. Kebenaran atau ketidak benaran suatu hipotesis tidak pernah diketahui secara pasti, kecuali jika kita memeriksa atau meneliti seluruh populasinya. Hal ini seringkali tidak mungkin dilakukan karena berbagai kendala, baik waktu, biaya, maupun tenaga yang harus dialokasikan untuk hal tersebut. Oleh karena itu, pemeriksaan terhadap kebenaran atau ketidak benaran hipotesis tersebut umumnya dilakukan melalui pengambilan sampel dari populasi.

Tujuan pengujian hipotesis adalah untuk memilih salah satu dari dua hipotesis tentang parameter populasi, yang keduanya saling bertentangan, yaitu hipotesis nol dinyatakan dengan  $H_0$ , dan hipotesis alternative atau hipotesis penelitian, dinyatakan dengan  $H_1$ . Kedua hipotesis tersebut bersifat saling asing (mutually exclusive), artinya jika satu hipotesis ditolak, maka sebagai konsekuensinya, hipotesis lain diterima.

Pengujian hipotesis berdasarkan pada konsep pembuktian melalui pengkontradikian, yaitu jika analisis

terhadap sampel menunjukkan ketidak konsistenan dengan hipotesis yang diuji, maka hipotesis tersebut ditolak dan disimpulkan bahwa hipotesis tersebut salah. Sebaliknya, jika analisis sampel tersebut konsisten dengan hipotesis yang diuji, maka hipotesis tersebut diterima. Namun demikian, penerimaan terhadap suatu hipotesis semata-mata sebagai akibat tidak cukupnya bukti untuk menolak hipotesis dan tidak berarti bahwa hipotesis tersebut adalah benar.

Prosedur pengujian hipotesis terdiri dari komponen komponen berikut ini:

- a. Hipotesis nol
- b. Hipotesis alternative atau hipotesis penelitian
- c. Statistic uji (metode analisis data)
- d. Daerah kritis atau daerah penolakan hipotesis

Hipotesis nol adalah asumsi atau anggapan yang berkaitan dengan nilai parameter populasi yang akan diuji. Hipotesis nol umumnya menyatakan bahwa nilai parameter populasi tersebut sama dengan suatu nilai tertentu. Misalkan  $A$  adalah parameter populasi, maka  $H_0$  umumnya dinyatakan dengan bentuk.

$H_0 : A = A_0$  dimana  $A_0$  adalah suatu nilai tertentu

hipotesis alternative merupakan suatu pernyataan alternative jika asumsi atau anggapan tentang parameter populasi tersebut ternyata salah atau ditolak

#### **D. Tugas**

1. Apa yang dimaksud dengan hipotesis statistic?
2. Jelaskan persamaan dan perbedaan antara konsep-konsep pendugaan parameter dan pengujian hipotesis
3. Sebutkan komponen-komponen prosedur pengujian hipotesis
4. Jelaskan apakah yang dimaksud dengan hipotesis nol? Dan apakah yang dimaksud dengan hipotesis alternative?

5. Pedoman apakah yang digunakan untuk menentukan suatu statistic uji?
6. Sebutkan tiga jenis formal pasangan hipotesis dan jelaskan konsekuensinya terhadap daerah kritis pengujian?
7. Jelaskan apa yang dimaksud dengan kesalahan pengujian jenis I dan kesalahan pengujian jenis II ?
8. Jelaskan peluang terjadinya kesalahan pengujian jenis I meningkat, dan bagaimanakah pengaruhnya terhadap peluang terjadinya kesalahan pengujian jenis II?
9. Jelaskan apa yang dimaksud dengan taraf nyata dan kuasa pengujian ?
10. Jelaskan apa yang dimaksud dengan daerah penolakan atau daerah kritis pengujian?

#### **E. Rujukan**

- Aridinanti, L. dkk. (2003). *Pengantar Metoda Statistika*. Surabaya : Jurusan Statistika FMIPA ITS
- Aunuddin. (1989). *Analisis Data*. Bogor: Institut Pertanian Bogor
- Rumiati, A.T. (1998). *Konsep Peluang dan Statistika dalam Rekayasa*. Surabaya : Jurusan Statistika FMIPA ITS
- Walpole, E. R. (1995). *Pengantar Statistik Edisi Ketiga*. Jakarta : PT. Gramedia Pustaka Utama

## **F. Bacaan yang Dianjurkan**

Aridinanti, L. dkk. (2003). *Pengantar Metoda Statistika*.

Surabaya : Jurusan Statistika FMIPA ITS

Aunuddin. (1989). *Analisis Data*. Bogor: Institut

Pertanian Bogor

Rumiati, A.T. (1998). *Konsep Peluang dan Statistika dalam*

*Rekayasa*. Surabaya : Jurusan Statistika FMIPA ITS

Walpole, E. R. (1995). *Pengantar Statistik Edisi Ketiga*. Jakarta

: PT. Gramedia Pustaka Utama

# BAB V

## PENGUJIAN DATA

### A. Pendahuluan

Pada Bab ini akan dibahas konsep uji normalitas data, uji normalitas data dan uji shapiro-wilks serta aplikasinya dalam persoalan ekonomi dan bisnis.

#### **Tujuan Pembelajaran :**

Setelah mempelajari materi pada Bab ini mahasiswa diharapkan memahami konsep uji normalitas data, uji normalitas data dan bias mengaplikasikan uji shapiro-wilks serta aplikasinya dalam persoalan ekonomi dan bisnis.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nila  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. menjelaskan konsep uji normalitas data
2. menerapkan uji shapiro-wilks

### B. Konsep Uji Normalitas Data

Pada pengujian hipotesis khususnya uji parametrik (menduga parameter), uji normalitas sangatlah penting diperlukan untuk mengetahui apakah distribusi dari data sampel yang kita gunakan memenuhi asumsi berdistribusi

normal. penyimpangan terhadap asumsi kenormalan tersebut dapat berakibat terhadap keabsahan dalam penarikan kesimpulan, karena statistik hitung yang kita gunakan diturunkan dari fungsi distribusi normal. oleh karena itu, sebelum kita melakukan uji hipotesis, kita harus menguji terlebih dahulu kenormalan dari distribusi data sampel yang akan digunakan.

Uji normalitas yang biasa digunakan adalah uji Liliefors, uji Kolmogorov-Smirnov, uji Anderson-Darling dan uji Shapiro-Wilks. Akan tetapi dalam buku ini akan dibahas adalah uji Shapiro-Wilks karena menyesuaikan yang disediakan didalam R-Commander

### 1. Uji Normalitas Data

Pengujian asumsi ini digunakan untuk mengetahui residual dari model mengikuti pola distribusi distribusi normal. Secara visual asumsi ini dapat dideteksi dengan menggunakan plot sebaran peluang normal yaitu melihat plot hubungan antara sisaan dengan normal bakunya. Jika hasil membentuk garis lurus maka asumsi kenormalan terpenuhi.

Untuk mengetahui kenormalan suatu data agar bisa digunakan pengujian sebagai berikut :

$H_0$  : data berdistribusi normal

$H_1$  : data tidak berdistribusi normal

Dimana harus diketahui nilai *P-value* dan  $\alpha$  terlebih dahulu.

Dari output data diketahui bahwa *P-value* lebih kecil dari  $\alpha$  ( $P\text{-value} < \alpha$ ) maka hipotesis nol ( $H_0$  : data berdistribusi normal ) ditolak, apabila dari output data diketahui bahwa *P value* lebih besar dari  $\alpha$  ( $P\text{-value} > \alpha$  ) maka hipotesis tidak nol ( $H_0$  : data tidak berdistribusi normal) gagal ditolak.

## 2. Uji Saphiro-Wilk

Data berasal dari sampel random berukuran  $n$  ( $X_1, X_2, X_3, \dots, X_n$ ) dengan fungsi distribusi  $F(x)$  yang tidak diketahui. Uji Statistik (T) yaitu :

$$T_3 = \frac{1}{D} * \left[ \sum_{i=1}^k a_i (X^{(n-i+1)} - X(i)) \right]^2$$

dengan denominator :  $D = \sum_{i=1}^n (X_i - \bar{X})^2$

Dapat juga disubstitusikan menjadi :

$$G = b_n + c_n * \ln \left( \frac{T_3 - d_n}{1 - T_s} \right)$$

Diperoleh nilai  $Z_{hit} = \frac{\sum G}{\sqrt{n}}$  yang akan dibandingkan dengan  $Z_{tabel}$  dari distribusi normal standar dengan  $\alpha$  tertentu (umumnya 5%). Tolak  $H_0$  apabila nilai  $Z_{hit} < Z_{tabel}$ .

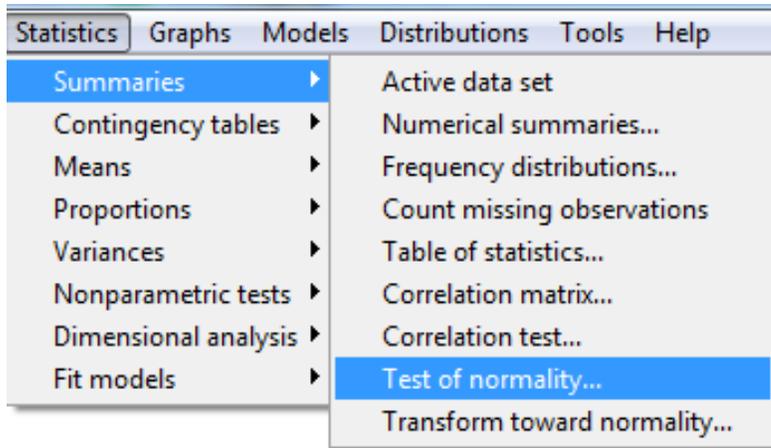
Sedangkan untuk uji Shapiro-Wil secara multivariate menggunakan statistik uji sebagai berikut :

$$W = (\sum a_t Z_t)^2 / \sum (z_t - \bar{z})^2$$

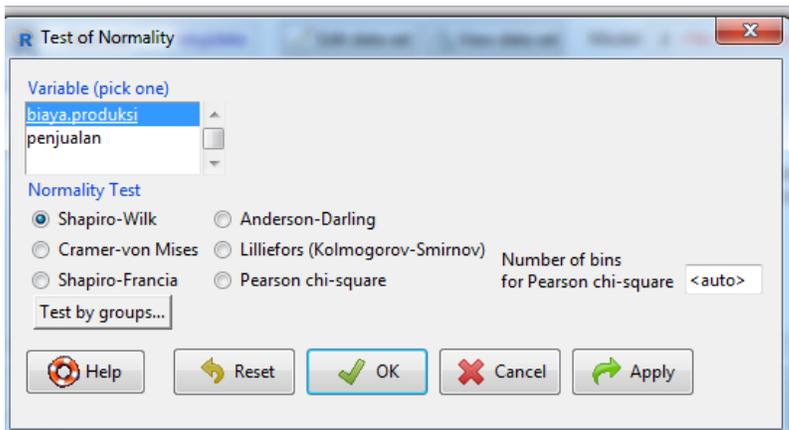
dimana  $z_t$  merupakan sampel yang telah diurutkan dari  $n$  pengamatan. Koefisien  $a_t$  bergantung pada matriks varians kovarians dari variabel random normal standar se-hingga lebih memudahkan pengambilan keputusan.

Misalkan  $z$  adalah kombinasi linear dari variabel yang didefinisikan  $x$ , maka sebagai penduga nilai minimum  $W$  memungkinkan untuk semua kombinasi linear. Yang lebih penting, uji ini melihat data dari arah nilai non-normality maksimum yang dinamakan statistik Shapiro-Wilk. Nilai kritis tidak tersedia untuk  $W$ , oleh karena itu perlu menghitung  $T_{bf}$ . atau menggunakan  $p$ -value dengan keterangan  $H_0$  Ditolak jika Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} < \alpha$ .

Contoh kasus akan dilakukan pengujian normalitas pada data tentang biaya produksi menggunakan R-Commander, adapun langkah-langkahnya adalah seperti berikut :



Setelah memilih menu tes of normality kemudian akan muncul menu tampilan seperti berikut :



Kemudian pilih variabel yang akan diuji normalitas pada datanya yaitu variabel biaya produksi, pilih normality test menggunakan metode Shapiro-Wilk. Hasil output perhitungan uji kenormalan shapiro-wilk pada R-Commander di peroleh :

```

> normalitytest (~biaya.produksi,
test="shapiro.test", data=latujidata)

shapiro-wilk normality test

data: biaya.produksi
w = 0.92663, p-value = 0.07266

```

Berdasarkan output tersebut dapat dibentuk Hipotesis yang digunakan dalam pengujian kenormalan Shapiro-Wilk adalah sebagai berikut :

Ho : data berdistribusi normal

H<sub>1</sub> : data tidak berdistribusi normal

dengan nilai tingkat kemaknaan  $\alpha$  sebesar 0,05 dengan nilai p-valuenya diperoleh sebesar 0,07266 sehingga dapat disimpulkan H<sub>0</sub> terima dikarekana nilai P-value lebih besar dari nilai kemaknaan sehingga kesimpulannya data untuk variable biaya penjualan tersebut adalah berdistribusi normal.

### C. Rangkuman

Pada pengujian hipotesis khususnya uji parametrik (menduga parameter), uji normalitas sangatlah penting diperlukan untuk mengetahui apakah distribusi dari data sampel yang kita gunakan memenuhi asumsi berdistribusi normal. penyimpangan terhadap asumsi kenormalan tersebut dapat berakibat terhadap keabsahan dalam penarikan kesimpulan, karena statistik hitung yang kita gunakan diturunkan dari fungsi distribusi normal. oleh karena itu, sebelum kita melakukan uji hipotesis, kita harus menguji terlebih dahulu kenormalan dari distribusi data sampel yang akan digunakan.

Untuk mengetahui kenormalan suatu data agar bisa digunakan pengujian sebagai berikut :

Ho : data berdistribusi normal

H<sub>1</sub> : data tidak berdistribusi normal

Dimana harus diketahui nilai *P-value* dan  $\alpha$  terlebih dahulu. Dari output data diketahui bahwa *P-value* lebih kecil dari  $\alpha$  ( $P\text{-value} < \alpha$ ) maka hipotesis nol ( $H_0$  : data berdistribusi normal ) ditolak, apabila dari output data diketahui bahwa *P-value* lebih besar dari  $\alpha$  ( $P\text{-value} > \alpha$ ) maka hipotesis tidak nol ( $H_0$  : data tidak berdistribusi normal) gagal ditolak. uji Shapiro-Wil secara multivariate menggunakan statistik uji sebagai berikut :

$$W = (\sum a_t Z_t)^2 / \sum (z_t - \bar{z})^2$$

dimana  $z_t$  merupakan sampel yang telah diurutkan dari  $n$  pengamatan. Koefisien  $a_t$  bergantung pada matriks varians kovarians dari variabel random normal standar se-hingga lebih memudahkan pengambilan keputusan.

Misalkan  $z$  adalah kombinasi linear dari variabel yang didefinisikan  $x$ , maka sebagai penduga nilai minimum  $W$  memungkinkan untuk semua kombinasi linear. Yang lebih penting, uji ini melihat data dari arah nilai non-normality maksimum yang dinamakan statistik Shapiro-Wilk. Nilai kritis tidak tersedia untuk  $W$ , oleh karena itu perlu menghitung  $T_{bf}$ . atau menggunakan *p-value* dengan keterangan  $H_0$  Ditolak jika Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} < \alpha$  .

#### **D. Tugas**

Lakukan pengujian normalitas pada data dari hasil pencarian yang sudah dilakukan menggunakan metode uji shapiro-wilks, kemudian berikan analisis dari hasil pengujian anda!

#### **E. Rujukan**

- Alvin C, Rencher. (1998). Multivariate Statistical Inference and Application. Canada: John Willey & Sons.Inc.
- Conover, W.J. (1980). Practical Nonparametric Statistics Second Edition. Canada: John Willey & Sons, Inc.

- Hair, dkk. (2006). *Multivariate Data Analysis sixth edition*.  
New Jersey: Pearson Plentice hall
- Zanbar, Achmad S. (2005). *Ilmu Statistika Pendekatan Teoritis dan Aplikatif disertai Contoh Penggunaan SPSS*.  
Bandung : Rekayasa Sains.
- Iriawan, Nur dan Septin Puji A. (2006). *Mengolah Data Statistik dengan Mudah Menggunakan Minitab 14*.  
Yogyakarta : Andi

#### **F. Bacaan yang Dianjurkan**

- Alvin C, Rencher. (1998). *Multivariate Statistical Inference and Application*. Canada: John Willey & Sons.Inc.
- Conover, W.J. (1980). *Practical Nonparametric Statistics Second Edition*. Canada: John Willey & Sons, Inc.
- Hair, dkk. (2006). *Multivariate Data Analysis sixth edition*.  
New Jersey: Pearson Plentice hall
- Zanbar, Achmad S. (2005). *Ilmu Statistika Pendekatan Teoritis dan Aplikatif disertai Contoh Penggunaan SPSS*.  
Bandung : Rekayasa Sains.
- Iriawan, Nur dan Septin Puji A. (2006). *Mengolah Data Statistik dengan Mudah Menggunakan Minitab 14*.  
Yogyakarta : Andi

# BAB VI

## UJI BEDA RATA-RATA

### A. Pendahuluan

Pada Bab ini akan dibahas tentang Uji T sampel tunggal, Uji T dua sampel berpasangan, Uji T dua sampel independen dan Anova (*Analysis Of Varian*), serta penerapannya dalam persoalan ekonomi dan bisnis.

#### **Tujuan Pembelajaran :**

Setelah mempelajari materi pada Bab ini mahasiswa diharapkan bias menerapkan Uji T sampel tunggal, Uji T dua sampel berpasangan, Uji T dua sampel independen dan Anova (*Analysis Of Varian*), serta penerapannya dalam persoalan ekonomi dan bisnis.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nilai  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan
3. CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
4. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Uji T sampel tunggal,
2. Uji T dua sampel berpasangan
3. Uji T dua sampel independen,
4. Anova (*Analysis Of Varian*),

## B. Pengujian Beda Rata-Rata

### 1. Uji T Sampel Tunggal

Seperti penaksiran selang kepercayaan, maka uji menyangkut satu rata-rata populasi  $\mu$  dengan  $\sigma^2$  tidak diketahui, tentunya melibatkan penggunaan distribusi t-Student. Peubah acak  $X_1, X_2, \dots, X_n$  menyatakan sample acak dari distribusi normal dengan  $\mu$  dan  $\sigma^2$  tidak diketahui. Maka peubah acak  $\sqrt{n}(\bar{X} - \mu)/S$  berdistribusi t-Student dengan derajat kebebasan  $n-1$ . Struktur uji itu sama dengan kasus  $\sigma$  diketahui kecuali bahwa nilai  $\sigma$  pada uji statistik diganti dengan taksiran  $S$  hasil perhitungan dan distribusi normal diganti dengan distribusi-t. jadi, untuk uji hipotesis dua sisi :

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

penolakan  $H_0$  pada taraf keberartian  $\alpha$  bila statistik t hasil perhitungan :

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

melampaui  $t_{\alpha/2, n-1}$  atau kurang dari  $-t_{\alpha/2, n-1}$ . Untuk  $H_1: \mu > \mu_0$ , penolakan diambil  $t > t_{\alpha, n-1}$ . Untuk  $H_1: \mu < \mu_0$ , daerah kritisnya adalah  $t < -t_{\alpha, n-1}$ . Atau apabila p-value  $< \alpha$  maka  $H_0$  ditolak.

Contoh kasus :

Seorang peneliti melakukan penelitian tentang penjualan suatu produk yang dilakukan oleh sales yang ada di perusahaannya. Dimana manajer penjualan menganggap ada satu sales yang penjualannya sebanyak 333 buah sehingga penjualan sales tersebut berbeda dengan rekan-rekannya yang lain. Benarkah pernyataan manajer tersebut?

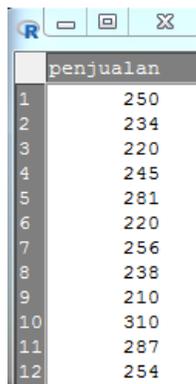
Berdasarkan kasus diatas maka metode yang tepat untuk digunakan yaitu uji t sampel tunggal karena digunakan untuk menguji rata-rata penjualan dari perusahaan yang dilakukan sales apakah penjualannya berbeda secara signifikan dengan rata-rata yang ditetapkan sebesar 333 buah.

Sehingga peneliti dapat membentuk hipotesisnya adalah sebagai berikut :

$H_0 = \mu = 333$  , jumlah produk yang dijual sales tersebut tidak berbeda dengan rata-rata dari produk yang terjual yaitu 333

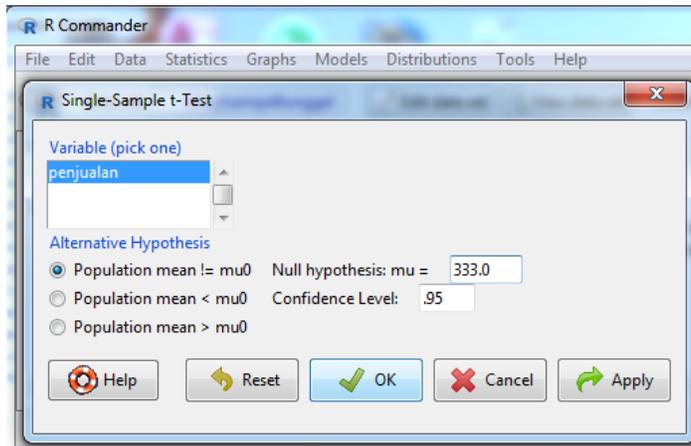
$H_1 = \mu \neq 333$  , jumlah produk yang dijual sales tersebut berbeda dengan rata-rata dari produk yang terjual yaitu 333

Prosedur melakukan pengujian rata-rata sampel tunggal menggunakan R-Commander adalah menginput data di Excel kemudian melakukan impor data menggunakan R-Commander dengan klik menu Data → impor data → from excel file...misalkan dengan memberi nama data yang akan kita gunakan tersebut adalah latujitsampeltunggal, pilih sheet yang digunakan dalam menginput data di excel. Untuk mengecek data yang kita input dapat dilihat di view data set. Adapun datanya adalah sebagai berikut :



	penjualan
1	250
2	234
3	220
4	245
5	281
6	220
7	256
8	238
9	210
10	310
11	287
12	254

Langkah selanjutnya yaitu melakukan uji t dengan sampel tunggal dengan memasukkan nilai rata-rata yang ditetapkan oleh peneliti yaitu  $\mu$  sebesar 333 dengan mengisikna confidence level 0,95 karena kita menggunakan  $\alpha = 0,05$  kemudian klik ok. Adapun langkahnya adalah seperti berikut:



Output yang diperoleh dari pengujian rata-rata sampel tunggal diatas yaitu:

```
> with(latujuitsampeltunggal, (t.test(penjualan,
alternative='two.sided',
mu=333.0,
conf.level=.95)))
```

One Sample t-test

```
data: penjualan
t = -9.582, df = 11, p-value = 0.000001131
alternative hypothesis: true mean is not equal to
333
95 percent confidence interval:
 231.4473 269.3860
sample estimates:
mean of x
 250.4167
```

Interpretasi dari hasil perhitungan Rcommander diatas diperoleh nilai statistic t sebesar -9,582 dengan p-value sebesar 0,000001131 untuk pengambilan keputusan paling mudah cukup melihat pada nilai p-value yang akan dibandingkan dengan tingkat kemaknaan yang digunakan dalam penelitian ini yaitu  $\alpha = 0,05$ . Sehingga jelas tampak hasil bahwa nilai p-value lebih kecil dari pada  $\alpha = 0,05$  dimana nilai probabilitasnya jauh dibawah  $\alpha = 0,05$ , maka kesimpulannya  $H_0$  ditolak maka dapat diputuskan bahwa penjualan sales tersebut memang berbeda disbanding dengan rata-rata penjualan yang lain.

## 2. Uji T Dua Sampel Berpasangan

Uji t dua sampel berpasangan (paired t test) digunakan untuk menguji dua sampel yang terdiri dari satu populasi yang sama akan tetapi diperlakukan dua perlakuan yang berbeda misanya dibedakan dengan Sebelum dan Sesudah perlakuan (*Before - After t test*)

Hipotesis statistika yang digunakan adalah

$$H_0 : \delta=0$$

$$H_1 : \delta \neq 0$$

dengan  $\delta$  adalah rata-rata (populasi) selisih nilai 1 dan 2 (pre dan post)

rumus

Statistik uji dalam pengujian ini adalah (Johnson dan Bhat tacharyya, 1996)

$$t = \frac{\bar{d} - d_o}{s / \sqrt{n}}$$

$d$  = rerata selisih nilai 1 dan 2 (pre dan post)

$s$  = simpangan baku selisih (beda) nilai

$n$  = ukuran atau besar sampel

Karena uji ini termasuk dalam uji dua arah dengan  $H_1$  bertanda tidak sama dengan maka  $H_0$  ditolak jika nilai uji  $t$  memenuhi daerah penolakan, yaitu

$$|t| > t_{\frac{\alpha}{2}; df=n-1} \text{ atau } p\text{-value} < \alpha .$$

Contoh kasusnya misalkan ada seorang mahasiswa akan melakukan penelitian tentang perbedaan harga saham antara sebelum dan sesudah adanya stock split pada perusahaan manufaktur yang melakukan stock split di Bursa Efek Indonesia. Peneliti menggunakan data sekunder yang diperoleh dari data-data perusahaan manufaktur di BEI dengan jumlah perusahaan sebanyak 12 perusahaan ( $N=12$ ). Lakukan pengujian hipotesis dengan tingkat kemaknaan  $\alpha = 0,05$  !

Untuk menyelesaikan pengujian hipotesis di atas metode yang cocok digunakan oleh peneliti adalah paired sampel t test (uji t dua sampel berpasangan) karena pada uji sampel ini dimaksudkan untuk mengetahui apakah ada perbedaan harga saham antara sebelum stock split dan sesudah stock split pada perusahaan di BEI.

Sehingga dapat dibentuk hipotesisnya dalam penelitian ini yaitu :

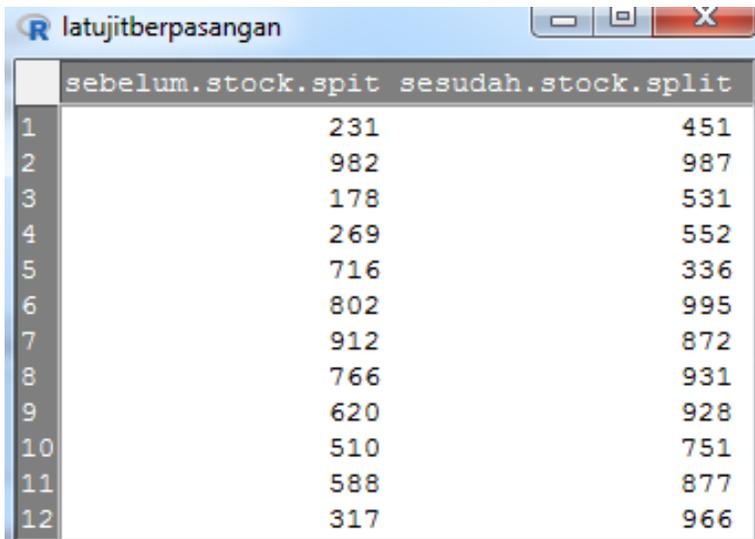
$H_0$  yaitu tidak ada perbedaan yang signifikan antara harga saham sebelum dan sesudah adanya stock split

$H_1$  yaitu ada perbedaan yang signifikan antara harga saham sebelum dan sesudah adanya stock split

Menyelesaikan pengujian tersebut menggunakan bantuan R-Commander dengan langkah-langkah sebagai berikut :

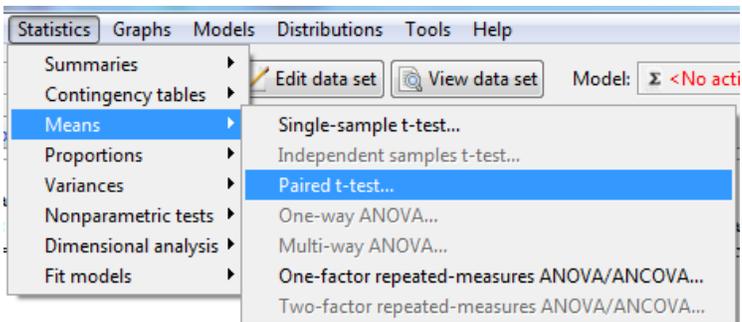
Input data di excel, kemudian lakukan import data pada R-Commander dengan menu sebagai berikut Data → import data → from Excel file...masukkan nama untuk data set misalkan latujitberpasangan pilih sheet mana

yang digunakan untuk menyimpan datanya di Excel, sehingga datanya dapat diperoleh yaitu dengan mengecek pada view data set. Adalah sebagai berikut :

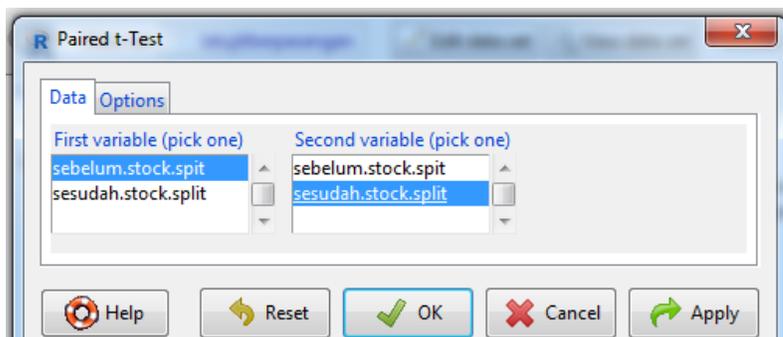


	sebelum.stock.spit	sesudah.stock.split
1	231	451
2	982	987
3	178	531
4	269	552
5	716	336
6	802	995
7	912	872
8	766	931
9	620	928
10	510	751
11	588	877
12	317	966

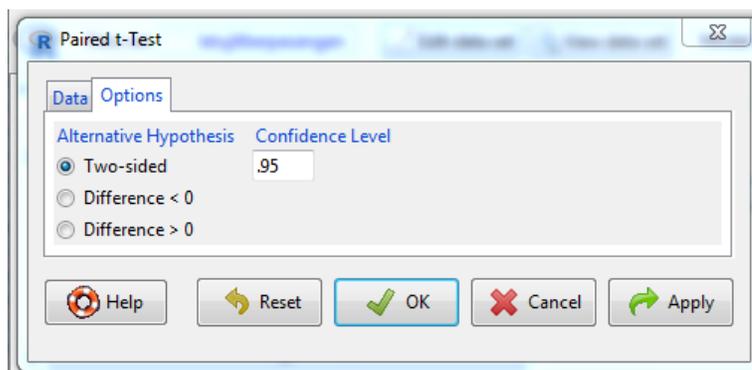
Setelah data sudah siap langkah selanjutnya yaitu melakukan analisis uji sampel berpasangan dengan langkah klik statistics → means → paired t test



Selanjutnya akan terbuka menu Paired t Test seperti berikut :



Klik option untuk memilih hipotesis alternatifnya dan memasukkan nilai confiden level yang digunakan, karena taraf signifikansi yang digunakan 0,05 maka konfiden level intervanya adalah 0,95 sehingga dimasukkan ke program .95 kemudian klik OK



Sehingga diperoleh output hasil pengujiannya perbedaan rata-rata sampel berpasangan yang diperoleh dari contoh kasus diatas adalah sebagai berikut :

```
> with(latujitberpasangan,  
(t.test(sebelum.stock.spit, sesudah.stock.split,  
alternative='two.sided', conf.level=.95,  
paired=TRUE)))
```

Paired t-test

```

data: sebelum.stock.split and sesudah.stock.split
t = -2.643, df = 11, p-value = 0.02287
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
  -349.14372  -31.85628
sample estimates:
mean of the differences
                -190.5

```

Dari output diatas diperoleh nilai statistic t hitung sebesar -2,643 dengan nilai p-value sebesar 0,02287 pengambilan keputusan paling mudah dilihat dengan membandingkan p-value dengan tingkat kemaknaan(  $\alpha$  ) yang digunakan dengan demikian nilai p-value lebih kecil dari tingkat kemaknaan (  $\alpha = 0,05$  ), sehingga dapat diambil kesimpulan bahwa  $H_0$  pengujian tersebut ditolak. Hal ini berarti terdapat perbedaan yang signifikan antara harga saham sebelum dan sesudah adanya stock split. Dapat diketahui bahwa nilai harga saham sesudah stock split lebih tinggi dari pada harga saham sebelum stock split. Dengan ini dapat disimpulkan pula bahwa adanya stock split akan meningkatkan harga saham.

### 3. Uji T Dua Sampel Independen

Dalam uji menyangkut dua rata-rata keadaan yang lebih umum berlaku ialah keadaan dengan varians tidak diketahui. Jika dianggap kedua distribusi normal dan bahwa  $\sigma_1 = \sigma_2 = \sigma$ , maka uji-t gabungan ( sering disebut uji-t dua sampel ) dapat digunakan. Uji statistik tersebut berbentuk :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_o}{S_p \sqrt{1/n_1 + 1/n_2}}$$

untuk :

$$S_p^2 = \frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

distribusi-t digunakan di sini dan bila hipotesisnya dua arah maka hipotesis tidak ditolak bila :

$$-t_{\alpha/2, n_1+n_2-2} < t < t_{\alpha/2, n_1+n_2-2}$$

tandingan satu sisi menimbulkan daerah kritis satu arah. Sebagai contoh, untuk  $H_0: \mu - \mu_0 > d_0$ , tolak  $H_0: \mu - \mu_0 = d_0$

bila  $t > t_{\alpha/2, n_1+n_2-2}$ . atau  $pvalue < \alpha$ . maka  $H_0$  ditolak.

Contoh kasunya misalkan :

Seorang peneliti akan membandingkan suatu produk merk A dan Merk B yang dibandingkan berdasarkan lama waktu penggunaannya sampai kerusakan. Untuk menduga beda kedua merk tersebut, dilakukan percobaan mangambil 8 sampel yang berbeda tapi tetap dari dua merk yang akan di uji untuk masing-masing merk. Untuk satu merk dipilih produk dan dipasang secara acak kemudian dicatat lama waktu penggunaannya sampai merk tersebut rusak. Peneliti ingin menguji apakah kedua merk tersebut memberikan lama waktu penggunaan yang sama dengan menggunakan tingkat kemaknaan  $\alpha = 0,05$ .

Penyelesaian :

Hipotesis penelitian tersebut adalah :

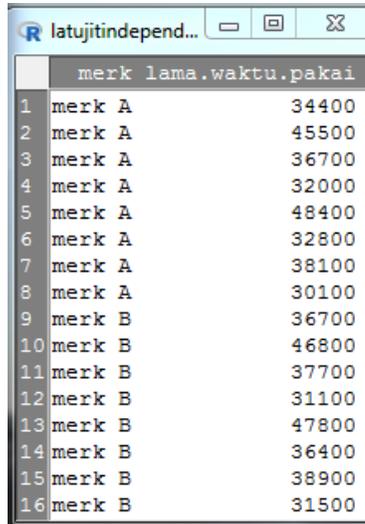
$H_0$  = tidak ada perbedaan yang signifikan lama pakai sampai rusak antara merk A dengan merk B

$H_1$  = terdapat perbedaan yang signifikan lama pakai sampai rusak antara merk A dengan merk B

Prosedur melakukan pengujian rata-rata dua sampel saling bebas menggunakan R-Commnander adalah sebagai berikut :

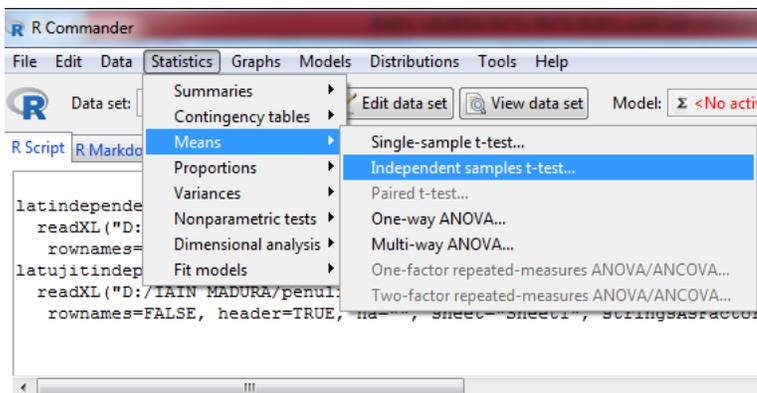
Input data kedalam Excell kemudian melakukan impor data menggunakan R-Commander dengan menu klik Data

→ impor data → from excel file...untuk melihat data yang kita input sesuai apa tidak dengan klik view data set missal data kita namai latujitindependen. Adapun datanya adalah sebagai berikut :



	merk	lama.pakai	waktu.pakai
1	merk A		34400
2	merk A		45500
3	merk A		36700
4	merk A		32000
5	merk A		48400
6	merk A		32800
7	merk A		38100
8	merk A		30100
9	merk B		36700
10	merk B		46800
11	merk B		37700
12	merk B		31100
13	merk B		47800
14	merk B		36400
15	merk B		38900
16	merk B		31500

Langkah selanjutnya yaitu melakukan uji t sampel independen dengan klik menu statistics→means→independent samples t-test...

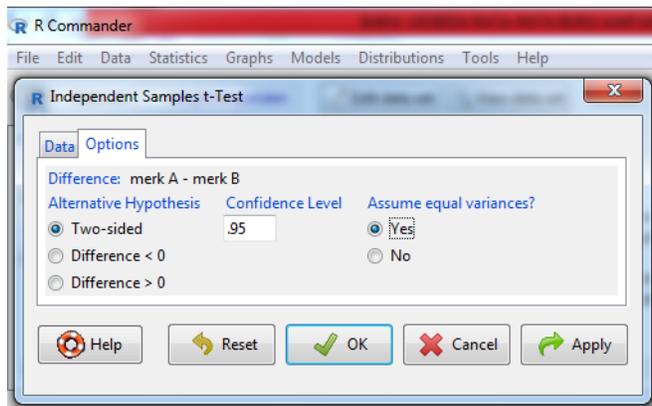


Pilih variabel pada kelompok dan variabel lama waktu pakai pada variabel respon.



Pilih hipotesis alternatif yang sesuai dengan kasus diatas, yaitu two sided (dua arah).

Masukkan interval keyakinan dengan 0,95 karena tingkat kemaknaan  $\alpha = 0,05$ , asumsi kesamaan variansi pilih ya. Kemudian klik OK.



Ouput hasil perhitungan menggunakan R-Commander pengujian beda rata-rata dua sampel independen untuk kasus ini diperoleh sebagai berikut :

```
> t.test(lama.waktu.pakai~merk,
alternative='two.sided',
conf.level=.95,
var.equal=TRUE, data=latujitindependen)
```

## Two Sample t-test

```
data: lama.waktu.pakai by merk
t = -0.34948, df = 14, p-value = 0.7319
alternative hypothesis: true difference in means
between group merk A and group merk B is not
equal to 0
95 percent confidence interval:
-7939.93 5714.93
sample estimates:
mean in group merk A mean in group merk B
37250.0 38362.5
```

Dari output hasil perhitungan Program R tersebut diperoleh nilai statistik t sebesar  $-0,34948$  dengan p-value sebesar  $0,7319$ . Untuk pengambilan keputusan langsung saja dengan mudah cukup melihat nilai dari p-value dibandingkan dengan nilai tingkat kemaknaan  $\alpha = 0,05$ . Sehingga dapat diperoleh bahwa nilai p-value lebih besar dari nilai  $\alpha = 0,05$ , sehingga keputusannya  $H_0$  pengujian tersebut diterima. Jadi kesimpulannya tidak ada perbedaan yang signifikan lama pakai sampai rusak antara merk A dengan merk B, berarti bahwa rata-rata lama pemakaian merk A dan Merk B sampai mengalami kerusakan adalah sama

#### 4. Anova (Analisis Of Varians)

Merupakan pengembangan dari uji t untuk dua sampel bebas. Bertujuan mengetahui ada tidaknya perbedaan nilai rerata lebih dari 2 kelompok

Sumber Variasi	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F <sub>STAT</sub>
perlakuan	k-1	JKP	KTP=JKP/(k-1)	KTP/KTE
error	n-k	JKE	KTE=JKE/(n-k)	
total	n-1	JK Total		

kriteria penolakan  $H_0$  :  $F_{STAT} > F_{k-1, n-k, 1-\alpha}$  atau  $p\text{-value} < \alpha$ .

Rumus untuk perhitungan nilai-nilai JKP, JKE dan JK Total yaitu :

Rumus Jumlah Kuadrat Perlakuan (JKP) :

$$JKP = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2 = \sum_{i=1}^k \frac{T_i^2}{n_i} - \frac{T^2}{n}$$

Rumus Jumlah Kuadrat Error (JKE) :

$$JKE = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

Rumus Jumlah Kuadrat Total (JKT) :

$$JK \text{ Total} = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 = \sum_{ij} x_{ij}^2 - \frac{T^2}{n}$$

Dengan keterangan

$T_i$  = total pengamatan populasi (perlakuan) ke  $-i$

$T$  = total pengamatan seluruhnya

$n_i$  = banyaknya pengamatan perlakuan ke  $-i$

$n$  = banyaknya pengamatan seluruhnya.

Contoh kasus :

Seorang peneliti ingin melakukan penelitian di pasar modal dimana peneliti tersebut beranggapan bahwa rata-rata penjualan saham perusahaan yang go public sama di tiga bursa efek A, B dan C dengan alternative tak sama. Untuk menguji anggapan itu, di setiap bursa efek akan diteliti sebanyak 5 perusahaan yang go public sebagai sampel acak. Dengan  $\alpha = 0,05$  ujilah pendapat tersebut.

Penyelesaian :

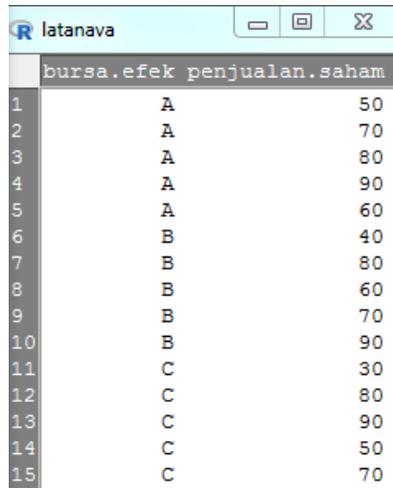
Hipotesis penelitian tersebut dapat dibentuk yaitu :

$H_0$  adalah  $\mu_1 = \mu_2 = \mu_3$ , artinya rata-rata penjualan saham perusahaan yang go public tersebut untuk masing-masing bursa efek adalah sama.

$H_1$  = sekarang kurangnya terdapat satu tanda tidak sama dengan artinya rata-rata penjualan saham perusahaan yang go public tersebut untuk masing-masing bursa efek adalah tidak sama.

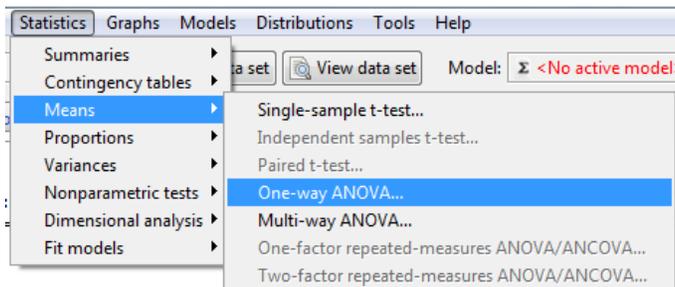
Prosedur melakukan pengujian analisis variansi (ANOVA) menggunakan R-Commander yaitu :

Input data kedalam Excell kemudian melakukan impor data menggunakan R-Commander dengan menu klik Data → impor data → from excel file...untuk melihat data yang kita input sesuai apa tidak dengan klik view data set missal data kita namai latanava. Adapun datanya adalah sebagai berikut

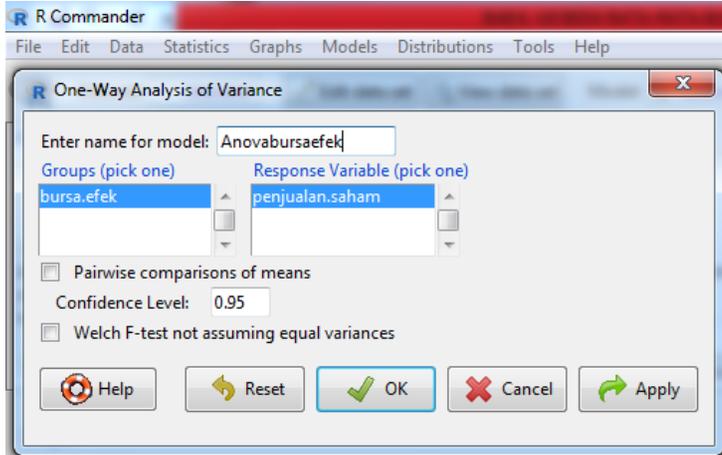


1	A	50
2	A	70
3	A	80
4	A	90
5	A	60
6	B	40
7	B	80
8	B	60
9	B	70
10	B	90
11	C	30
12	C	80
13	C	90
14	C	50
15	C	70

Langkah selanjutnya yaitu melakukan uji ANOVA dengan klik menu statistics → means → One-way ANOVA...



Masukkan nama model yang akan dibuat misalnya Anovabursaefek. Pilih variabel bursa efek menjadi grupnya dan variabel penjualan saham sebagai variabel respon. Langsung klik OK



Ouput hasil perhitungan menggunakan R-Commander pengujian ANOVA dalam kasus ini tepatnya Anava satu arah karena hanya terdapat satu grup variabel saja untuk kasus ini diperoleh sebagai berikut :

```
> Anovabursaefek <- aov(penjualan.saham ~
bursa.efek, data=latanava)

> summary(Anovabursaefek)
              Df Sum Sq Mean Sq F value Pr(>F)
bursa.efek    2     93    46.7    0.117  0.891
Residuals   12   4800   400.0

> with(latanava, numSummary(penjualan.saham,
groups=bursa.efek, statistics=c("mean", "sd")))
      mean      sd data:n
A    70 15.81139      5
B    68 19.23538      5
C    64 24.08319      5
```

Dari output program R diatas diperoleh bahwa nilai statistic F sebesar 0,117 dengan p-value 0,891. Dengan

demikian p-value lebih besar dari nilai tingkat kemaknaan  $\alpha = 0,05$ , sehingga dapat diambil keputusan bahwa  $H_0$  uji tersebut diterima. Hal ini berarti rata-rata penjualan saham perusahaan yang go public tersebut untuk masing-masing bursa efek adalah sama. Artinya tidak terjadi perbedaan yang signifikan mean antara ketiga bursa efek tersebut hal ini dapat dilihat pada nilai mean yang diperoleh yaitu bursa efek A sebesar 70, mean untuk bursa efek B = 68 dan rata-rata untuk bursa efek C adalah 64 jadi selisih rata-rata antar grup bursa efek tidak terlalu signifikan selisihnya. Sehingga dapat disimpulkan rata-rata penjualan saham perusahaan yang go public tersebut untuk masing-masing bursa efek adalah sama.

### C. Rangkuman

Uji beda rata-rata digunakan untuk menguji apakah mean(rata-rata) terdapat perbedaan yang signifikan atau tidak antara satu sampel, dua sampel atau lebih dari dua sampel. Adapun metode pengujiannya ada empat metode yaitu : Uji T sampel tunggal, Uji T dua sampel berpasangan, Uji T dua sampel independen dan Anova (*Analysis Of Variance*). uji t sampel tunggal menyangkut satu rata-rata populasi  $\mu$  dengan  $\sigma^2$  tidak diketahui, tentunya melibatkan penggunaan distribusi t-Student. Peubah acak  $X_1, X_2, \dots, X_n$  menyatakan sample acak dari distribusi normal dengan  $\mu$  dan  $\sigma^2$  tidak diketahui. Maka peubah acak  $\sqrt{n}(\bar{X} - \mu)/S$  berdistribusi t-Student dengan derajat kebebasan  $n-1$ . Struktur uji itu sama dengan kasus  $\sigma$  diketahui kecuali bahwa nilai  $\sigma$  pada uji statistik diganti dengan taksiran  $S$  hasil perhitungan dan distribusi normal diganti dengan distribusi-t. jadi, untuk uji hipotesis dua sisi :

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

penolakan  $H_0$  pada taraf keberartian  $\alpha$  bila statistik  $t$  hasil perhitungan :

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

melampaui  $t_{\alpha/2, n-1}$  atau kurang dari  $-t_{\alpha/2, n-1}$ . Untuk  $H_1: \mu > \mu_0$ , penolakan diambil  $t > t_{\alpha, n-1}$ . Untuk  $H_1: \mu < \mu_0$ , daerah kritisnya adalah  $t < -t_{\alpha, n-1}$ . Atau apabila  $p\text{-value} < \alpha$  maka  $H_0$  ditolak.

Uji  $t$  dua sampel berpasangan (paired  $t$  test) digunakan untuk menguji dua sampel yang terdiri dari satu populasi yang sama akan tetapi diperlakukan dua perlakuan yang berbeda misanya dibedakan dengan Sebelum dan Sesudah perlakuan (*Before - After  $t$  test*)

Hipotesis statistika yang digunakan adalah

$$H_0 : \delta = 0$$

$$H_1 : \delta \neq 0$$

dengan  $\delta$  adalah rata-rata (populasi) selisih nilai 1 dan 2 (pre dan post)

rumus

Statistik uji dalam pengujian ini adalah (Johnson dan Bhattacharyya, 1996)

$$t = \frac{\bar{d} - d_0}{s / \sqrt{n}}$$

$\bar{d}$  = rerata selisih nilai 1 dan 2 (pre dan post)

$s$  = simpangan baku selisih (beda) nilai

$n$  = ukuran atau besar sampel

karena uji ini termasuk dalam uji dua arah dengan  $H_1$  bertanda tidak sama dengan maka  $H_0$  ditolak jika nilai uji  $t$

memenuhi daerah penolakan, yaitu  $|t| > t_{\frac{\alpha}{2}; df=n-1}$  atau  $p\text{-value} < \alpha$ .

Uji menyangkut dua rata-rata keadaan yang lebih umum berlaku ialah keadaan dengan varians tidak diketahui. Jika dianggap kedua distribusi normal dan bahwa  $\sigma_1 = \sigma_2 = \sigma$ , maka uji-t gabungan ( sering disebut uji-t dua sampel ) dapat digunakan. Uji statistik tersebut berbentuk :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{S_p \sqrt{1/n_1 + 1/n_2}}$$

untuk :

$$S_p^2 = \frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

distribusi-t digunakan di sini dan bila hipotesisnya dua arah maka hipotesis tidak ditolak bila :

$$-t_{\alpha/2, n_1+n_2-2} < t < t_{\alpha/2, n_1+n_2-2}$$

tandingan satu sisi menimbulkan daerah kritis satu arah. Sebagai contoh, untuk  $H_0: \mu - \mu_0 > d_0$ , tolak  $H_0: \mu - \mu_0 = d_0$  bila  $t > t_{\alpha/2, n_1+n_2-2}$ . atau  $pvalue < \alpha$ . maka  $H_0$  ditolak.

Merupakan pengembangan dari uji t untuk dua sampel bebas. Bertujuan mengetahui ada tidaknya perbedaan nilai rerata lebih dari 2 kelompok, kriteria penolakan  $H_0 : F_{STAT} > F_{k-1, n-k, 1-\alpha}$  atau  $p-value < \alpha$  ..

#### D. Tugas

Lakukan pengujian uji beda rata-rata menggunakan metode Uji T sampel tunggal, Uji T dua sampel berpasangan, Uji T dua sampel independen dan Anova (Analisis Of Varian) dari hasil pencarian data yang sudah dilakukan, kemudian berikan analisis dari hasil pengujian anda.

#### E. Rujukan

Soepeno, Bambang. (1997). Statistik Terapan dalam Penelitian Ilmu-ilmu Sosial dan Pendidikan. Indonesia : PT Rineka Cipta

Sudjana. (1996). Metoda Statistika. Bandung : Tarsito  
Sugiyono. (2003). Statistika untuk Penelitian. Bandung : CV  
Alfabeta  
Zanten, V. W. (1980). Statistik untuk Ilmu- ilmu Sosial.  
Jakarta : Gramedia

**F. Bacaan yang Dianjurkan**

Soepeno, Bambang. (1997). Statistik Terapan dalam  
Penelitian Ilmu-ilmu Sosial  
dan Pendidikan. Indonesia : PT Rineka Cipta  
Sudjana. (1996). Metoda Statistika. Bandung : Tarsito  
Sugiyono. (2003). Statistika untuk Penelitian. Bandung: CV  
Alfabeta  
Zanten, V. W. (1980). Statistik untuk Ilmu- ilmu Sosial. Jakarta:  
Gramedia

# BAB VII

## PENGUJIAN HIPOTESIS TENTANG PROPORSI

### A. Pendahuluan

Pada Bab ini akan dibahas tentang Pengujian Hipotesis Tentang Proporsi yang terdiri dari uji proporsi sampel tunggal dan uji proporsi dua sampel beserta prakteknya dalam persoalan ekonomi dan bisnis.

#### **Tujuan pembelajaran :**

Setelah mempelajari materi pada bab ini mahasiswa diharapkan mampu menerapkan Pengujian Hipotesis Tentang Proporsi yang terdiri dari uji proporsi Dari Satu sampel dan Selisih Proporsi Dari Dua sampel beserta prakteknya dalam persoalan ekonomi dan bisnis.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nilai  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Pengujian Hipotesis Untuk Proporsi Dari Satu Populasi
2. Pengujian Hipotesis Untuk Selisih Proporsi Dari Dua sampel

## B. Uji Proporsi

Dalam pengujian hipotesis proporsi untuk sampel terdiri dari dua jenis yaitu pengujian proporsi sampel tunggal dan pengujian proporsi dua sampel. Pengujian dilakukan untuk mengetahui apakah proporsi yang dihipotesiskan didukung informasi dari data sampel (apakah proporsi sampel berbeda dengan proporsi yang dihipotesiskan) dan untuk pengujian proporsi dua sampel dimana Pengujian proporsi digunakan ketika akan membandingkan apakah proporsi pada sampel pertama lebih kecil, sama atau lebih besar dibandingkan proporsi pada sampel kedua.

### 1. Pengujian Hipotesis Untuk Proporsi Dari Satu sampel

Dalam pengujian proporsi untuk sampel tunggal metode yang digunakan adalah uji Chi kuadrat untuk satu sampel, karena teknik staitstik ini digunakan untuk menguji hipotesis deskriptif bila dalam populasi terdiri atas dua atau lebih klas, data berbentuk nominal atau ordinal dan sampelnya besar. Yang dimaksud hipotesis deskriptif disini bisa merupakan estimasi/dugaan terhadap ada tidaknya perbedaan frekuensi antara kategori satu dan kategori lain dalam sebuah sampel tentang sesuatu hal/

Rumus Statistik uji yang dipakai adalah :

$$\chi^2 = \sum_{i=1}^k \frac{(f_o - f_h)^2}{f_n}$$

Dimana

$f_o$  adalah frekuensi yang diobservasi

$f_h$  adalah frekuensi yang diharapkan

Kaidah keputusannya yaitu  $H_0$  Ditolak jika nilai

$$\chi^2 > \chi_{\alpha, dk=1}^2$$

Atau jika Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} < \alpha$ .

Berikut ini akan diberikan contoh kasus uji proporsi sampel tunggal untuk menguji hipotesis satu sampel yang terdiri atas dua kategori.

Contoh kasus :

Salah satu perusahaan parfum ingin mengeluarkan suatu produknya. Seorang manajer produksi ingin melakukan pengujian untuk mengetahui apakah produk tersebut akan berpeluang sama untuk dibeli antara pembeli wanita dan pembeli pria. Untuk itu maka diperlukan penelitian. Populasi penelitian adalah masyarakat perkotaan. Sampel diperoleh 200 orang pembeli pria dan 100 orang pembeli wanita.

Hipotesis yang diajukan adalah

$H_0$  adalah frekuensi/jumlah masyarakat yang membeli parfum antara banyaknya pria dan wanita adalah sama (peluang produk parfum untuk disenangi antara pria dan wanita adalah sama) .

$H_1$  adalah frekuensi/ jumlah masyarakat yang membeli membeli parfum antara banyaknya pria dan wanita adalah tidak sama.pembeli pria dan wanita berpeluang berbeda untuk membeli parfum).

Dalam hal ini

$$H_0 : P_1 = P_2 = 0,5$$

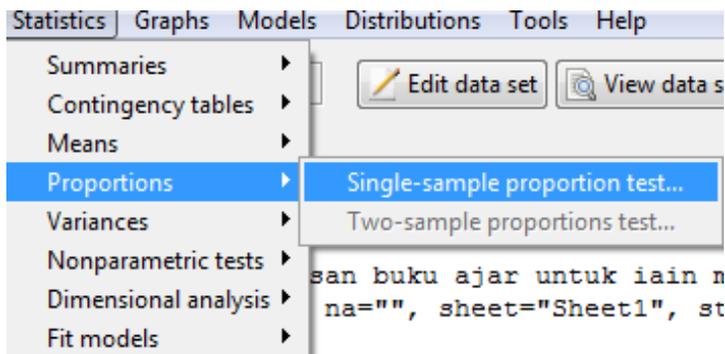
$$H_1 : P_1 \neq P_2 \neq 0,5$$

Perhitungan akan dilakukan menggunakan Program R khususnya dengan paket R-Commander yaitu sebagai berikut :

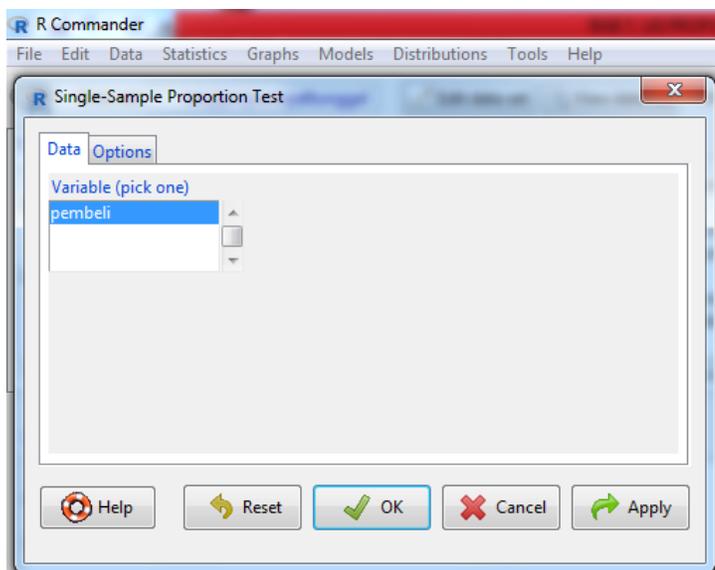
Input data ke excel kemudian impor data dengan menu Data → impor data → from excel file... kita namai datanya latujipropsampeltunggal.

Kemudian dilakukan langkah uji proporsi untuk sampel tunggal dengan langkah

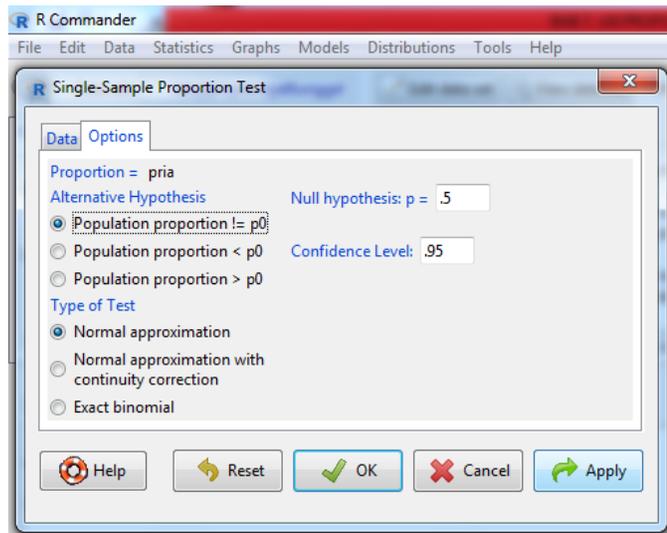
statistics → Proportions → Single-sample proportion test...



Pilihan hanya ada satu variable pembeli saja karena sampel tunggal yang akan diuji proporsinya, sehingga muncul menu seperti berikut :



Kemudian klik option untuk memilih berapa proporsi yang akan diujikan pada hipotesis null nya dana menentukan nila confiden intervalnya dan untuk memilih alternative hipotesis untuk populasi proporsinya dan tipe dari pengujiannya. kemudian klik OK.



Sehingga diperoleh output sebagai berikut :

```
> local({
+   .Table <- xtabs(~ pembeli , data=
latujipropsampeltunggal )
+   cat("\nFrequency counts (test is for first
level):\n")
+   print(.Table)
+   prop.test(rbind(.Table),
alternative='two.sided', p=.5, conf.level=.95,
correct=FALSE)
+ })
```

Frequency counts (test is for first level):

```
pembeli
  pria wanita
  200   100
```

1-sample proportions test without continuity correction

```
data: rbind(.Table), null probability 0.5
X-squared = 33.333, df = 1, p-value =
0.000000007764
alternative hypothesis: true p is not equal to
0.5
95 percent confidence interval:
```

```
0.6115125 0.7176066
sample estimates:
      p
0.6666667
```

Berdasarkan output diatas dapat dijelaskan bahwa banyak pria yang membeli parfum adalah 200 orang dan wanita 100 orang. Untuk melakukan pengambilan keputusan atau kesimpulan dapat dilihat dari nilai chi kuadratnya ternyata diperoleh nilai sebesar 33,333 dimana nilai tersebut akan dibandingkan dengan nilai chi kuadrat tabel berdasarkan derajat kebebasan yang sesuai sehingga diperoleh nilai chi kuadrat tabelnya sebesar 3,841 karena nilainya chi kuadrat hitung lebih besar dari tabel ( $33,333 > 3,841$ ) sesuai dengan ketentuan maka  $H_0$  ditolak dan  $H_1$  diterima.. Atau dengan mudah dalam pengambilan keputusan kita cukup melihat p-value diperoleh nilainya sebesar 0.00000007764 yang akan dibandingkan dengan tingkat kemaknaan yang akan digunakan dalam penelitian yaitu  $\alpha = 0,05$  dengan ketentuan apabila nilainya yaitu  $p\text{-value} < \alpha$ ,  $H_0$  Ditolak jadi kesimpulannya kaum pria cenderung membeli produk parfum tersebut.

## 2. Pengujian Hipotesis Untuk Selisih Proporsi Dari Dua sampel

Pengujian proporsi dari dua sampel digunakan untuk menguji signifikansi proporsi dua sampel yang tidak berpasangan. Dilakukan apabila data hasil pengamatan tidak dapat diukur karena skala pengukurannya dalam bentuk nominal atau ordinal. Hipotesis yang akan diuji biasanya menyatakan bahwa dua kelompok kasus berbeda dalam satu karakteristik tertentu. Untuk menguji hipotesis ini, kita hitung jumlah individu dari tiap kelompok yang termasuk ke dalam

berbagai kategori dan membandingkan jumlah individu dari kelompok dalam berbagai kategori dengan kelompok yang lain.

Rumus Statistik uji yang dipakai adalah :

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Dimana

$O_{ij}$  adalah jumlah pengamatan pada baris ke-i dan kolom ke-j

$E_{ij}$  adalah nilai harapan pengamatan pada baris ke-i dan kolom ke-j

Kaidah keputusannya yaitu  $H_0$  Ditolak jika nilai  $\chi^2 > \chi^2_{\alpha, dk=1}$  atau jika Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} < \alpha$ .

Contoh kasus :

Suatu penelitian ingin mengetahui apakah ada perbedaan proporsi antara yang menjawab ya dan tidak dalam melakukan hubungan kerja sama kembali dimasa akan datang dalam kedua kelompok konsumen industri yaitu industri majalah dan industri koran. Dengan menggunakan  $\alpha = 0,05$ .

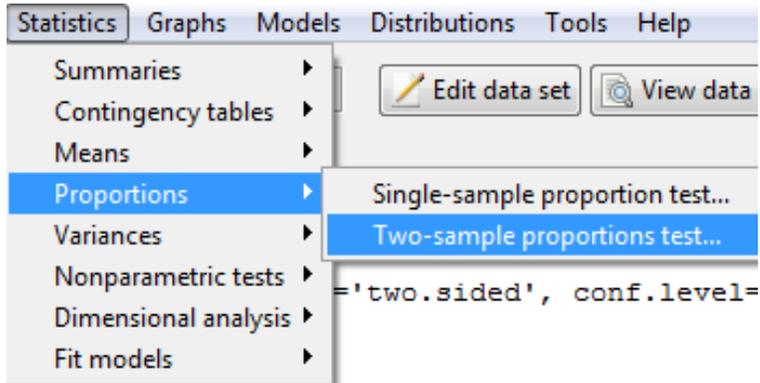
Berdasarkan kasus diatas dapat dibentuk hipotesisnya adalah sebagai berikut :

$H_0$  :  $P_1 = P_2$  atau proporsi yang menjawab tidak mau melakukan hubungan kerjasama kembali dimasa akan datang pada konsumen perusahaan majalah dan koran adalah sama

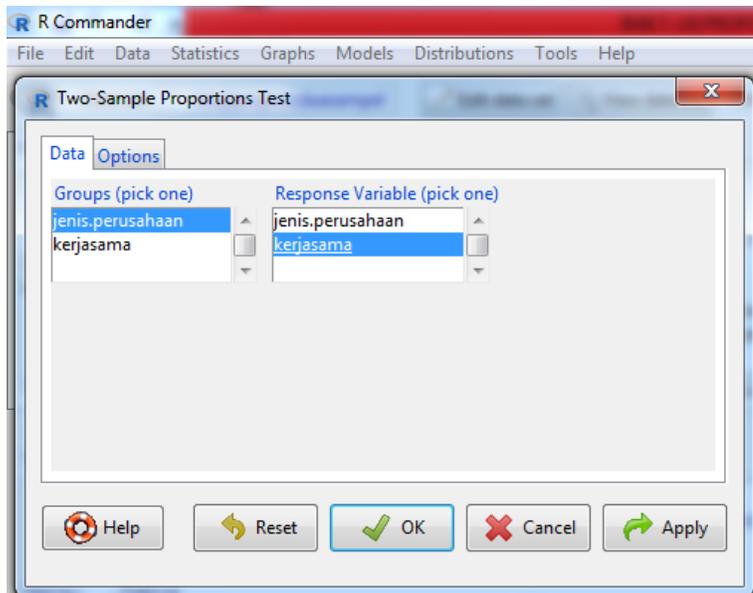
$H_1$  :  $P_1 \neq P_2$  atau terdapat perbedaan proporsi yang menjawab tidak mau melakukan hubungan kerjasama

kembali dimasa akan datang pada konsumen perusahaan majalah dan koran.

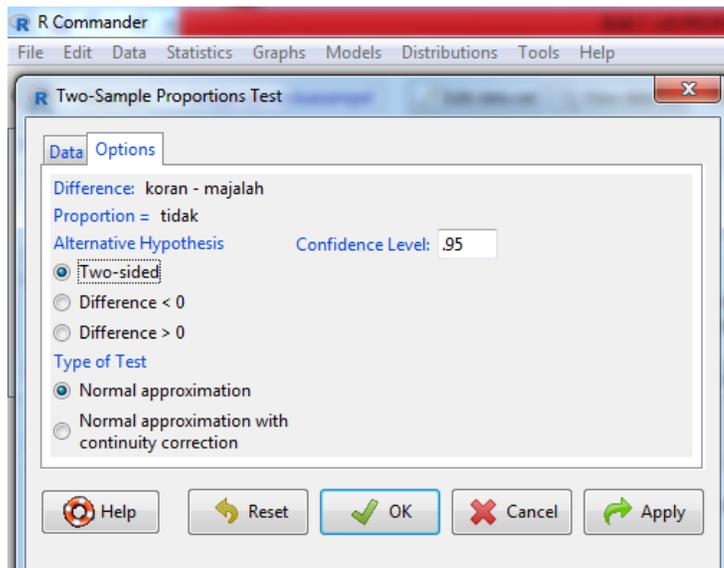
Untuk menyelesaikan pengujian proporsi kasus diatas digunakanlah bantuan R-Commander dengan dimulai menginput data hasil surveynya ke dalam excell kemudian melakukan impor data dengan menu Data → impor data → from excel file... kita namai datanya latuji proporsiduasampel, kemudian siap melakukan uji proporsi dua arah seperti menu tampilan berikut :



Hasil dari langkah tersebut yaitu sebagai berikut :



Lakukan dengan memilih pada tampilan data dengan menjadikan variable jenis perusahaan sebagai groups dan variable kerjasama sebagai variable respon. Jangan lupa pilih option untuk menentukan setelan dengan pilih hipotesis alternatifnya yaitu dua arah dengan tingkat kepercayaan 0,95 karena peneliti menggunakan  $\alpha = 0,05$  dengan tipe pengujiannya. Kemudian klik OK.



Sehingga diperoleh tampilan output dari perhitungan diatas adalah sebagai berikut :

```
> local({
+   .Table <-
+   xtabs(~jenis.perusahaan+kerjasama,
+   data=latujiproporsiduasampel)
+   cat("\nPercentage table:\n")
+   print(rowPercents(.Table))
+   prop.test(.Table, alternative='two.sided',
+   conf.level=.95, correct=FALSE)
+ })
Percentage table:
               kerjasama
jenis.perusahaan tidak ya Total Count
```

```

koran      52.1 47.9  100   48
majalah   57.7 42.3  100   52
2-sample test for equality of proportions
without continuity correction
data:      .Table
X-squared = 0.31728, df = 1, p-value = 0.5732
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.2510366  0.1388571
sample estimates:
   prop 1    prop 2
0.5208333 0.5769231

```

Berdasarkan hasil output di atas dapat dijelaskan bahwa hasil pengujiannya yaitu  $H_0$  diterima artinya proporsi yang menjawab tidak mau melakukan hubungan kerjasama kembali dimasa akan datang untuk konsumen perusahaan majalah dan koran adalah sama. Hal ini ditunjukkan oleh nilai p-value diperoleh sebesar 0,5732 dimana nilai tersebut jelas lebih besar dari nilai tingkat kemaknaan yang digunakan yaitu  $\alpha = 0,05$ . Sehingga dapat disimpulkan bahwasannya jenis industri media baik majalah atau Koran tidak memberikan perbedaan yang signifikan terhadap keinginan suatu konsumen untuk akan melakukan jalinan kerjasama kembali dengan membentuk hubungan dengan perusahaan di waktu yang akan datang.

### C. Rangkuman

Dalam pengujian hipotesis proporsi untuk sampel terdiri dari dua jenis yaitu pengujian proporsi sampel tunggal dan pengujian proporsi dua sampel. Pengujian dilakukan untuk mengetahui apakah proporsi yang dihipotesiskan didukung informasi dari data sampel ( apakah proporsi sampel berbeda dengan proporsi yang dihipotesiskan) dan untuk pengujian proporsi dua sampel dimana Pengujian

proporsi digunakan ketika akan membandingkan apakah proporsi pada sampel pertama lebih kecil, sama atau lebih besar dibandingkan proporsi pada sampel kedua.

Dalam pengujian proporsi untuk sampel tunggal metode yang digunakan adalah uji Chi kuadrat untuk satu sampel, karena teknik staitstik ini digunakan untuk menguji hipotesis deskriptif bila dalam populasi terdiri atas dua atau lebih klas, data berbentuk nominal atau ordinal dan sampelnya besar. Yang dimaksud hipotesis deskriptif disini bisa merupakan estimasi/dugaan terhadap ada tidaknya perbedaan frekuensi antara kategori satu dan kategori lain dalam sebuah sampel tentang sesuatu hal/

Rumus Statistik uji yang dipakai adalah :

$$\chi^2 = \sum_{i=1}^k \frac{(f_o - f_h)^2}{f_n}$$

Dimana

$f_o$  adalah frekuensi yang diobservasi

$f_h$  adalah frekuensi yang diharapkan

Kaidah keputusannya yaitu  $H_0$  Ditolak jika nilai  $\chi^2 > \chi_{\alpha, dk=1}^2$  atau jika Nilai tingkat kemaknaan yang diperoleh yaitu *p-value*  $< \alpha$  .

Pengujian proporsi dari dua sampel digunakan untuk menguji signifikansi proporsi dua sampel yang tidak berpasangan. Dilakukan apabila data hasil pengamatan tidak dapat diukur karena skala pengukurannya dalam bentuk nominal atau ordinal. Hipotesis yang akan diuji biasanya menyatakan bahwa dua kelompok kasus berbeda dalam satu karakteristik tertentu. Untuk menguji hipotesis ini, kita hitung jumlah individu dari tiap kelompok yang termasuk ke dalam berbagai kategori dan membandingkan jumlah individu dari kelompokdalam berbagai kategori dengan kelompok yang lain.

Rumus Statistik uji yang dipakai adalah :

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Dimana

$O_{ij}$  adalah jumlah pengamatan pada baris ke-i dan kolom ke-j

$E_{ij}$  adalah nilai harapan pengamatan pada baris ke-i dan kolom ke-j

Kaidah keputusannya yaitu  $H_0$  Ditolak jika nilai  $\chi^2 > \chi^2_{\alpha, dk=1}$  atau jika Nilai tingkat kemaknaan yang diperoleh yaitu *p-value*  $< \alpha$ .

#### **D. Tugas**

Lakukan pengujian uji proporsi sampel tunggal dan uji proporsi dua sampel dari hasil pencarian data yang sudah dilakukan, kemudian berikan analisis dari hasil pengujian anda.

#### **E. Rujukan**

- Cononer, W. J. (1999). Practical Nonparametric Statistic. Thirt Edition. Texas, America : John Wiley & Sons, INC
- Daniel, Wayne, W. (1989). Statistik Nonparametrik Terapan. Jakarta : PT Gramedia
- Siegel, Sidney. (1986). Statistik Nonparametrik untuk Ilmu-ilmu Sosial. Jakarta : Gramedia
- Sugiyono.( 2003). Statistik Nonparametrik untuk Penelitian. Bandung : CV Alfabeta

## **F. Bacaan yang Dianjurkan**

- Conover, W. J. (1999). Practical Nonparametric Statistics. Third Edition. Texas, America : John Wiley & Sons, INC
- Daniel, Wayne, W. (1989). Statistik Nonparametrik Terapan. Jakarta : PT Gramedia
- Siegel, Sidney. (1986). Statistik Nonparametrik untuk Ilmu-ilmu Sosial. Jakarta : Gramedia
- Sugiyono. (2003). Statistik Nonparametrik untuk Penelitian. Bandung : CV Alfabeta

# BAB VIII

## TABEL KONTINGENSI

### A. Pendahuluan

Pada Bab ini akan dibahas hal-hal yang mendasar berkenaan dengan tabel kontingensi khususnya tabel kontingensi dua arah dan tabel kontingensi dua arah menggunakan data frekuensi.

#### **Tujuan pembelajaran :**

Setelah mempelajari materi pada bab ini mahasiswa diharapkan menerapkan tabel kontingensi khususnya tabel kontingensi dua arah dan tabel kontingensi dua arah menggunakan data frekuensi.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nilai  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM) ). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Menerapkan tabel kontingensi dua arah
2. Menerapkan tabel kontingensi dua arah menggunakan data frekuensi

### B. Konsep Tabel Kontingensi

Tabel kontingensi biasanya disebut dengan tabel silang (*crosstab*) adalah sebuah tabel silang yang terdiri atas satu

baris atau lebih, dan satu kolom atau lebih. Ciri penggunaan crosstab adalah data input yang berskala nominal atau ordinal, seperti tabulasi antara gender seseorang dengan tingkat pendidikan orang tersebut, pekerjaan seseorang dengan sikap orang tersebut dengan suatu produk tertentu, dan lainnya. Sebenarnya pada data metrik (interval atau rasio) secara prinsip bisa juga dilakukan *crosstab*. Hanya pada data metrik, ada kemungkinan data mempunyai desimal, seperti panjang 1,25 meter, panjang 1,26 meter, panjang 1,27 meter, dan seterusnya. Semuanya mempunyai nilai berbeda sehingga harus dibuatkan banyak kolom; maka bias terjadi jumlah baris atau kolom menjadi demikian banyak dan tidak efektif untuk mendeskripsikan data. Untuk itulah pembuatan crosstab data metrik biasanya dilihat 'isi' datanya terlebih dahulu.

### 1. Tabel Kontingensi Dua Arah

Dalam praktek, pembuatan crosstab dapat juga disertai dengan penghitungan tingkat keeratan hubungan (asosiasi) antar isi crosstab. Alat statistik yang sering digunakan untuk mengukur asosiasi pada sebuah crosstab adalah chisquare. Prosedur uji Chi-

square Pearson atau  $\chi^2$

pada output diatas (untuk evaluasi dependensi antara dua perubah nonmetrik, skala nominal atau ordinal) adalah sebagai berikut

pembentukan hipotesisnya yaitu

H<sub>0</sub> : variable pada baris dan kolom saling bebas

H<sub>1</sub> : variable pada baris dan kolom saling terikat

$$\chi^2 = \sum_{i=1}^b \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Rumus statistic ujinya adalah

Dengan  $O_{ij}$  adalah jumlah pengamatan pada bari ke-i dan kolom ke-j,

$E_{ij}$  adalah jumlah harapan kejadian pada bari ke- $i$  dan kolom ke- $j$ ,

Perhitungan nilai harapan yaitu 
$$E_{ij} = \frac{n_i \cdot n_j}{n} .$$

Dengan keterangan  $n_i$  adalah total pengamatan baris ke- $i$

$n_j$  adalah total pengamatan kolom ke- $j$  dan  $n$  adalah total pengamatan keseluruhan.

Daerah Kriteria pengujian hipotesis Tolak  $H_0$  yang berarti kedua variable saling dependen (terikat) jika

$$\chi^2 > \chi_{\alpha, dk=(b-1)(k-1)}^2 \text{ atau nilai } p < \alpha ,$$

Dengan  $b$  adalah jumlah bari dan  $k$  adalah jumlah kolom

Contoh kasus :

Penelitian dilakukan untuk mengetahui adakah perbedaan antara tingkat pendidikan masyarakat dengan jenis bank yang dipilih untuk menyimpan uangnya. Pendidikan masyarakat dikelompokkan menjadi dua, yaitu lulusan SLTA dan Perguruan Tinggi. Sampel pertama sebanyak 80 lulusan SLTA dan sampel kedua sebanyak 70 lulusan Perguruan Tinggi. Berdasarkan angket yang diberikan kepada sampel yang lulusan SLTA, maka dari 80 orang itu yang memilih bank pemerintah sebanyak 60 orang, dan bank swasta sebanyak 20 orang. Selanjutnya dari kelompok sampel lulusan perguruan Tinggi, dari 70 orang itu yang memilih bank pemerintah sebanyak 30 orang dan bank swasta sebanyak 40 orang.

Hipotesis dalam penelitian tersebut dapat dibuat yaitu :

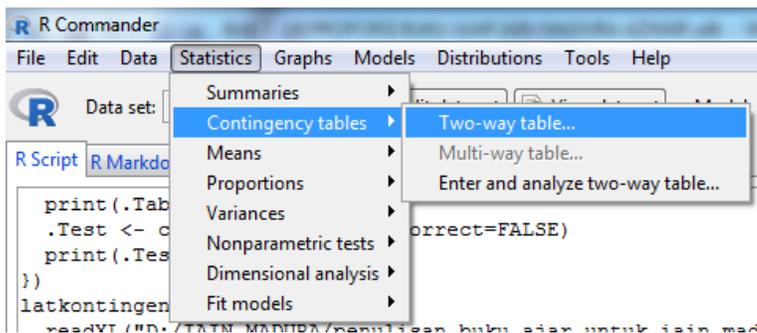
$H_0$  berbunyi tidak terdapat perbedaan tingkat pendidikan masyarakat dalam memilih dua jenis bank (atau tidak ada perbedaan antara jenjang pendidikan dengan jenis Bank yang dipilih).

$H_1$  berbunyi terdapat perbedaan tingkat pendidikan masyarakat dalam memilih dua jenis bank (atau ada perbedaan antara jenjang pendidikan dengan jenis Bank yang dipilih).

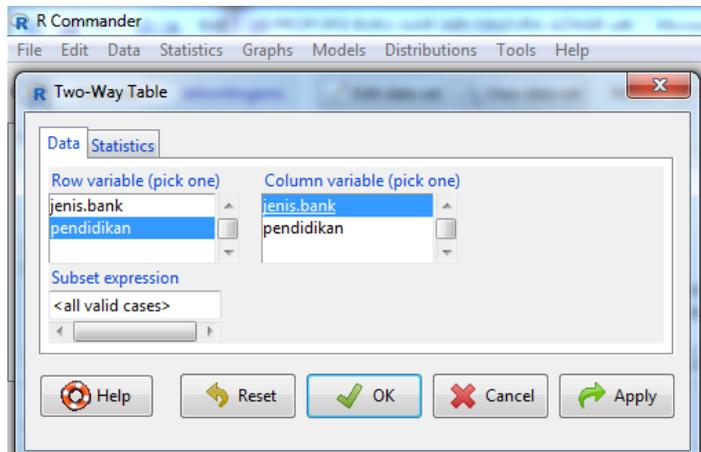
Menyelesaikan uji hipotesis tersebut menggunakan bantuan R-Commander dengan menginput terlebih dahulu data ke Excell kemudian melakukan impor data ke program R dengan menu Data → impor data → from excel file... kita namai datanya latkontingensi.

Kemudian dilakukan langkah uji tabel kontingensi untuk dua arah dengan langkah

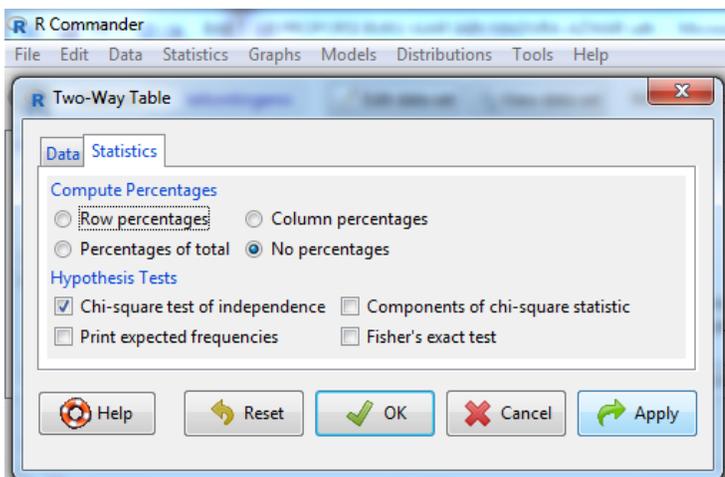
statistics → contingency tables → Two-way table...



Pilih variable pendidikan sebagai row variable dan jenis bank sebagai column variable, sehingga muncul menu seperti berikut :



Kemudian klik option untuk memilih uji hipotesisnya yang akan digunakan yaitu uji chi kuadrat dan untuk perhitungan prosentasenya klik no prosentages. kemudian klik OK.



Hasil dari output perhitungan menggunakan R-Comander adalah sebagai berikut :

```

> local({
+   .Table <- xtabs(~pendidikan+jenis.bank,
+     data=latkontingensi)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   .Test <- chisq.test(.Table, correct=FALSE)
+   print(.Test)
+ })
Frequency table:
                jenis.bank
pendidikan      bank pemerintah bank swasta
  lulusan PT                60             20
  lulusan SLTA              30             40
  Pearson's Chi-squared test
data:  .Table
X-squared = 16.071, df = 1, p-value = 0.000061

```

Berdasarkan dari hasil output diatas dapat dijelaskan sesuai kondisi yang telah dilakukan pada surveynya yaitu lulusan PT yang memilih bank pemerintah sebanyak 60 orang dan bank swasta sebanyak 20 orang dapat dilihat dari output tabel kontingensinya dan yang lulusan SLTA sebanyak 30 orang memilih bank pemerintah dan 40 memilih bank swasta. Sedangkan hasil pengujian hipotesisnya diperoleh nilai Chi kuadrat sebesar 16,071 dan p-value sebesar 0,000061 dari nilai tersebut jelas bahwa nilainya lebih kecil dari tingkat kemaknaan  $\alpha = 0,05$  maka  $H_0$  ditolak jadi dapat disimpulkan terdapat perbedaan tingkat pendidikan dalam memilih bank, dimana lulusan SLTA cenderung memilih Bank Pemerintah dan lulusan perguruan tinggi cenderung memilih bank Swasta. Sehingga dapat ditarik saran Bank Swasta perlu mempromosikan ke masyarakat yang berpendidikan Tinggi dan sebaliknya Bank Pemerintah perlu promosi ke masyarakat yang berpendidikan SLTA.

## 2. Tabel Kontingensi Langsung Input Menggunakan Data Frekuensi

Chi kuadrat k-sampel digunakan untuk menguji hipotesis komparatif lebih dari dua sampel bila datanya berbentuk nominal dan sampelnya besar. cara perhitungan dapat menggunakan rumus yang telah ada, atau dapat menggunakan tabel kontingensi r x k (r baris kali k kolom). pada bab ini di berikan contoh penggunaan tabel kontingensi untuk perhitungan Chi kuadrat karena lebih mudah.

kategori	Populasi/sampel				jumlah
	1	2...	j...	k	
1	$n_{11} (e_{11})$	$n_{12} (e_{12})$	$n_{1j} (e_{1j})$	$n_{1k} (e_{1k})$	$n_{1.}$
2	$n_{21} (e_{21})$	$n_{22} (e_{22})$	$n_{2j} (e_{2j})$	$n_{2k} (e_{2k})$	$n_{2.}$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
i	$n_{i1} (e_{i1})$	$n_{i2} (e_{i2})$	$n_{ij} (e_{ij})$	$n_{ik} (e_{ik})$	$n_{i.}$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
r	$n_{r1} (e_{r1})$	$n_{r2} (e_{r2})$	$n_{rj} (e_{rj})$	$n_{rk} (e_{rk})$	$n_{r.}$
jumlah	$n_{.1}$	$n_{.2}$	$n_{.j}$	$n_{.k}$	$n = n_{...}$

n adalah jumlah sampel.

Rumus Statistik uji yang dipakai adalah :

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Dimana

$O_{ij}$  adalah jumlah pengamatan pada baris ke-i dan kolom ke-j

$E_{ij}$  adalah nilai harapan pengamatan pada baris ke- $i$  dan kolom ke- $j$

Rumus Statistik uji yang dipakai adalah :

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Dimana

$O_{ij}$  adalah jumlah pengamatan pada baris ke- $i$  dan kolom ke- $j$

$E_{ij}$  adalah nilai harapan pengamatan pada baris ke- $i$  dan kolom ke- $j$

Dengan perhitungan untuk nilai harapan ini adalah sebagai berikut :

$$E_{ij} = \frac{n_{i.} n_{.j}}{n}$$

Dengan kriteria pengujian hipotesis terima  $H_0$  apabila nilai Chi kuadrat hitung lebih kecil dari nilai chi kuadrat tabel dengan derajat kebebasan yaitu banyak kolom - 1

kali banyak baris -1  $\chi^2 > \chi^2_{\alpha, dk=(b-1)(k-1)}$  dan taraf signifikansi tertentu. Atau nilai  $p > \alpha$ ,

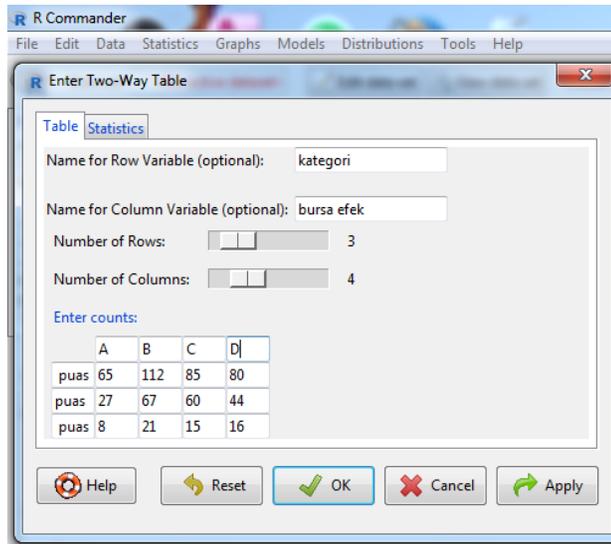
Contoh kasus :

Seorang peneliti akan membuktikan bahwa ada pendapat yang mengatakan bahwa tak ada perbedaan perbedaan dari pembeli saham yang sangat puas, puas dan tak puas dari empat penjualan saham dengan alternative ada perbedaan.

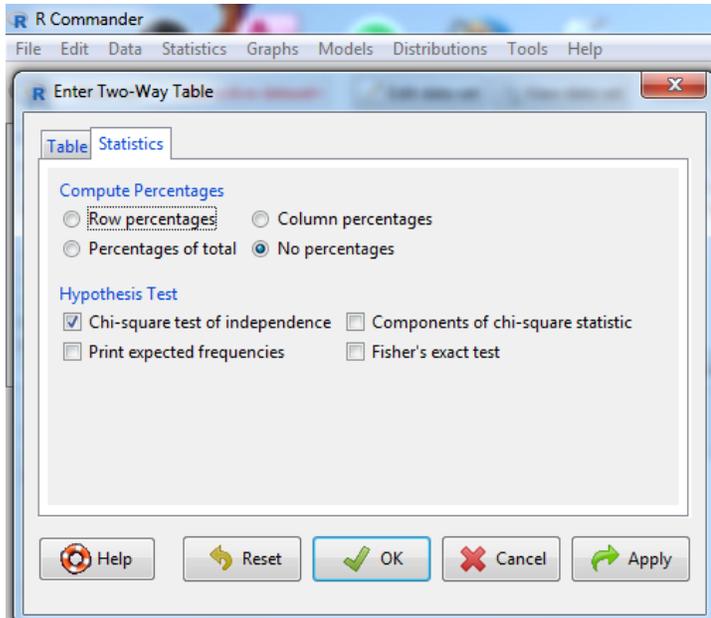
Sehingga dapat ditetapkan hipotesisnya untuk  $H_0$  berbunyi variable pada baris dan kolom independen. Dan  $H_1$  yaitu variable pada baris dan kolom dependen

Dari hasil penelitian sampel acak sebanyak 600 orang pembeli saham, dimana data dalam penelitian langsung

bias diinput ke dalam R-Commander menggunakan entry data frekuensi sehingga diperoleh data seperti berikut :



Dari kasus tersebut terdapat 4 kolom menyatakan jenis bursa efek dan 3 baris menyatakan tingkat kepuasan pembeli saham serta masukkan nilai frekuensinya. Untuk menghitungnya pilih statistics seperti berikut :



Pilih menu statistics kemudian centang sesuai keinginan peneliti misalkan pilih no percentages karena tidak ingin menampilkan nilai prosentase pilih uji hopotesis yang sesuai yaitu chi-square test of independen kemudian klik OK.

```
> .Table <- matrix(c(65,112,85,80,27,67,60,44,8,21,15,16),
  3, 4, byrow=TRUE)
> dimnames(.Table) <- list("kategori"=c("sangat
  puas", "puas", "tidak puas"), "bursa
  efek"=c("A", "B", "C", "D"))
> .Table # Counts
      bursa efek
kategori  A  B  C  D
sangat puas 65 112 85 80
puas        27 67 60 44
tidak puas  8 21 15 16

> .Test <- chisq.test(.Table, correct=FALSE)
> .Test
Pearson's Chi-squared test
data: .Table
X-squared = 4.5707, df = 6, p-value = 0.5999
```

Berdasarkan dari hasil output diatas menunjukkan bahwa nilai hasil uji Chi kuadrat adalah sebesar 4,5707 dan nilai p-value sebesar 0,5999. Sehingga dapat ditarik kesimpulan bahwa pembeli saham yang berpendapat sangat puas, puas dan tidak puas memiliki kesamaan persepsi terhadap bursa efek yang ada yaitu bursa efek A,B,C dan D.

### C. Rangkuman

Tabel kontingensi biasanya disebut dengan tabel silang (*crosstab*) adalah sebuah tabel silang yang terdiri atas satu baris atau lebih, dan satu kolom atau lebih. Ciri penggunaan crosstab adalah data input yang berskala nominal atau ordinal, seperti tabulasi antara gender seseorang dengan

tingkat pendidikan orang tersebut, pekerjaan seseorang dengan sikap orang tersebut dengan suatu produk tertentu, dan lainnya.

Dalam praktek, pembuatan crosstab dapat juga disertai dengan penghitungan tingkat keeratan hubungan (asosiasi) antar isi crosstab. Alat statistik yang sering digunakan untuk mengukur asosiasi pada sebuah crosstab adalah chisquare.

Prosedur uji Chi-square Pearson atau  $\chi^2$  pada output diatas (untuk evaluasi dependensi antara dua peubah nonmetrik, skala nominal atau ordinal) adalah sebagai berikut

pembentukan hipotesisnya yaitu

$H_0$  : variable pada baris dan kolom saling bebas

$H_1$  : variable pada baris dan kolom saling terikat

$$\chi^2 = \sum_{i=1}^b \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Rumus statistic ujinya adalah

Dengan  $O_{ij}$  adalah jumlah pengamatan pada bari ke-i dan kolom ke-j,

$E_{ij}$  adalah jumlah harapan kejadian pada bari ke-i dan kolom ke-j,

$$E_{ij} = \frac{n_i \cdot n_j}{n}$$

Perhitungan nilai harapan yaitu

Dengan keterangan  $n_i$  adalah total pengamatan baris ke-i

$n_j$  adalah total pengamatan kolom ke-j dan  $n$  adalah total pengamatan keseluruhan.

Daerah Kriteria pengujian hipotesis Tolak  $H_0$  yang berarti kedua variable saling dependen (terikat) jika

$\chi^2 > \chi^2_{\alpha, dk=(b-1)(k-1)}$  atau nilai  $p < \alpha$ , dengan b adalah jumlah bari dan k adalah jumlah kolom

Chi kuadrat k-sampel digunakan untuk menguji hipotesis komparatif lebih dari dua sampel bila datanya berbentuk nominal dan sampelnya besar. cara perhitungan dapat menggunakan rumus yang telah ada, atau dapat menggunakan tabel kontingensi  $r \times k$  ( $r$  baris kali  $k$  kolom). pada bab ini di berikan contoh penggunaan tabel kontingensi untuk perhitungan Chi kuadrat karena lebih mudah.

$n$  adalah jumlah sampel.

Rumus Statistik uji yang dipakai adalah :

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Dimana

$O_{ij}$  adalah jumlah pengamatan pada baris ke- $i$  dan kolom ke- $j$

$E_{ij}$  adalah nilai harapan pengamatan pada baris ke- $i$  dan kolom ke- $j$

Rumus Statistik uji yang dipakai adalah :

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Dimana

$O_{ij}$  adalah jumlah pengamatan pada baris ke- $i$  dan kolom ke- $j$

$E_{ij}$  adalah nilai harapan pengamatan pada baris ke- $i$  dan kolom ke- $j$

Dengan perhitungan untuk nilai harapan ini adalah sebagai berikut :

$$E_{ij} = \frac{n_i \cdot n_j}{n}$$

Dengan kriteria pengujian hipotesis terima  $H_0$  apabila nilai Chi kuadrat hitung lebih kecil dari nilai chi kuadrat tabel dengan derajat kebebasan yaitu banyak kolom - 1 kali banyak baris -1  $\chi^2 > \chi^2_{\alpha, dk=(b-1)(k-1)}$  dan taraf signifikansi tertentu. Atau nilai  $p > \alpha$ .

#### **D. Tugas**

Lakukan pengujian tabel kontingensi dua arah dan tabel kontingensi dengan entri data frekuensi dari hasil pencarian data yang sudah dilakukan, kemudian berikan analisis dari hasil pengujian anda.

#### **E. Rujukan**

- Cononer, W. J. (1999). Practical Nonparametric Statistic. Thirt Edition. Texas, America : John Wiley & Sons, INC
- Daniel, Wayne, W. (1989). Statistik Nonparametrik Terapan. Jakarta : PT Gramedia
- Siegel, Sidney. (1986). Statistik Nonparametrik untuk Ilmu-ilmu Sosial. Jakarta : Gramedia
- Sugiyono. (2003). Statistik Nonparametrik untuk Penelitian. Bandung : CV Alfabeta

#### **F. Bacaan yang Dianjurkan**

- Cononer, W. J. (1999). Practical Nonparametric Statistic. Thirt Edition. Texas, America : John Wiley & Sons, INC
- Daniel, Wayne, W. (1989). Statistik Nonparametrik Terapan. Jakarta : PT Gramedia
- Siegel, Sidney. (1986). Statistik Nonparametrik untuk Ilmu-ilmu Sosial. Jakarta : Gramedia
- Sugiyono. (2003). Statistik Nonparametrik untuk Penelitian. Bandung : CV Alfabeta

# BAB IX

## PENGUJIAN INSTRUMEN PENELITIAN

### A. Pendahuluan

Pada Bab ini akan dibahas uji validitas dan uji reliabilitas instrument penelitian dalam hal ini adalah kuesioner yang digunakan dalam penelitian, serta aplikasinya dalam persoalan ekonomi dan bisnis menggunakan bantuan program R.

#### **Tujuan Pembelajaran :**

Setelah mempelajari materi pada bab ini mahasiswa diharapkan memahami uji validitas dan reliabilitas instrument penelitian dan bisa menyelesaikan menggunakan program R.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nila  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Uji validitas
2. Uji reliabilitas

## **B. Konsep Pengukuran Validitas dan Reliabilitas**

Salah satu masalah utama dalam kegiatan penelitian adalah memperoleh data yang akurat dan objektif. Hal ini dikarenakan kesimpulan penelitian hanya dapat dipercaya apabila didasarkan pada informasi data yang juga dapat dipercaya. Tetapi, informasi data yang akurat dan objektif dalam penelitian biasanya tidak mudah diperoleh, terutama dalam penelitian yang menggunakan alat ukur berupa kuesioner. Selain konsep mengenai variabel yang diukur tidak selalu mudah dioperasionalkan, prosedur pengukuran juga menjadi penentu apakah informasi data yang dihasilkan dapat dipercaya.

Setiap alat ukur berupa kuesioner mempunyai kriteria untuk dapat dinyatakan sebagai alat ukur kuesioner yang baik, yakni mampu memberikan informasi berupa data yang dapat dipercaya yaitu meliputi reliabel, valid, standar, ekonomis dan praktis. Sifat reliabel dan valid diperlihatkan oleh tingginya reliabilitas dan validitas hasil ukur suatu kuesioner. Suatu instrumen ukur yang tidak reliabel atau tidak valid akan memberikan informasi yang tidak akurat mengenai keadaan individu atau subjek. Apabila informasi yang salah itu dengan sengaja digunakan sebagai dasar pertimbangan dalam mengambil keputusan maka kesimpulan dan keputusan itu merupakan kesimpulan atau keputusan yang tidak tepat. Dalam hal inilah pentingnya masalah konsep pengukuran validitas dan reliabilitas.

Validitas berasal dari kata *validity* yang mempunyai arti sejauhmana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Suatu instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila alat ukur tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukan pengukuran tersebut. Jadi pengertian validitas sangat erat kaitannya

dengan masalah tujuan pengukuran. Oleh karena itu, tidak ada yang berlaku umum untuk semua tujuan pengukuran. Suatu alat ukur biasanya hanya merupakan ukuran yang valid untuk suatu tujuan yang spesifik. Berdasarkan cara estimasinya pada umumnya tipe validitas digolongkan dalam tiga golongan, yaitu :

- a. *content validity* (validitas isi) merupakan validitas yang diestimasi berdasarkan atas pengujian isi alat ukur berdasarkan analisis rasional yang logik dalam menjelaskan representasi dari ciri-ciri variabel yang hendak diukur;
- b. *construct validity* (validitas konstruk) merupakan tipe validitas yang diestimasi berdasarkan sejauhmana suatu konstruk teoritik atau *trait* yang hendak diukurnya dalam hal ini berupa kisi-kisi dari konsep alat ukur yang akan diajukan;
- c. *criterion-related validity* (validitas berdasarkan kriteria) merupakan tipe validitas yang diestimasi berdasarkan hasil nilai empirik dari koefisien korelasi antara skor sebagai prediktor dan skor suatu kriteria yang relevan.

Reliabilitas merupakan terjemahan dari kata *reliability* yang mempunyai asal kata *rely* dan *ability*. Pengukuran yang memiliki reliabilitas tinggi disebut sebagai pengukuran yang reliabel (*reliable*). Nama lain pada reliabilitas antara lain keterpercayaan, keterandalan, keajegan, kestabilan, konsistensi, dan sebagainya, namun pada prinsipnya konsep reliabilitas adalah sejauhmana hasil suatu pengukuran dapat dipercaya.

Pengertian reliabilitas alat ukur dan reliabilitas hasil ukur biasanya dianggap sama, tetapi ada hal yang perlu diperhatikan. Konsep reliabilitas alat ukur berkaitan erat dengan masalah kesalahan pengukuran (*error of measurement*) yang menunjukkan pada sejauhmana

ketidakkonsistenan hasil pengukuran dilakukan ulang pada kelompok subjek yang sama. Sedangkan konsep reliabilitas hasil ukur berkaitan erat dengan kesalahan dalam pengambilan sampel (*sampling error*) yang mengacu kepada ketidakkonsistenan hasil ukur apabila pengukuran dilakukan ulang pada kelompok individu yang berbeda.

Pendekatan konsistensi internal dalam estimasi reliabilitas adalah untuk menghindari masalah-masalah yang biasanya ditimbulkan oleh pendekatan tes ulang dan oleh pendekatan bentuk paralel. Dalam pendekatan konsistensi internal prosedurnya hanya memerlukan satu kali pengenaan sebuah tes kepada sekelompok individu subjek (*single trial administration*).

### **1. Uji Validitas**

Konsep pengukuran validitas. Validitas berasal dari kata *validity* yang mempunyai arti sejauhmana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Suatu tes atau instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukan pengukuran tersebut. Jadi pengertian validitas sangat erat berkaitan dengan masalah tujuan pengukuran. Oleh karena itu, tidak ada yang berlaku umum untuk semua tujuan pengukuran. Suatu alat ukur biasanya hanya merupakan ukuran yang valid untuk suatu tujuan yang spesifik.

Istilah validitas empirik menunjukkan pada pengertian bahwa estimasi validitas termaksud dinyatakan oleh satu angka atau suatu koefisien, atau yang analisisnya dilakukan terhadap data yang diperoleh secara empirik, yaitu dari skor sekelompok subjek yang

dikenai tes tersebut. Ada dua macam validitas empirik, yaitu validitas konstruk dan validitas berdasar kriteria.

a. Validitas Konstruk

Sebagaimana telah dikemukakan, pengujian validitas konstruk sangat penting artinya terutama dalam pengembangan dan evaluasi terhadap skala-skala kepribadian. Prosedur pengujian validitas konstruk berawal dari hasil komputasi interkorelasi diantara berbagai hasil tes dan kemudian diikuti oleh analisis lebih lanjut terhadap matriks korelasi yang diperoleh, melalui berbagai metode.

Dua diantara pendekatan yang banyak dilakukan dalam pengujian validitas konstruk antara lain adalah pendekatan multitrait-multimethod dan pendekatan faktor analisis. Adapun penjelasannya adalah sebagai berikut :

b. Validitas Multitrait-Multimethod

Menurut Campbell dan fiske telah mengembangkan suatu pendekatan terhadap validitas konstruk yang mereka namakan pendekatan multitrait-multimethod. Pendekatan ini dapat digunakan bilamana terdapat dua trait atau lebih yang diukur oleh dua macam metode atau lebih.

Suatu contoh pendekatan ini dikemukakan oleh Allen dan Yen dengan mengandaikan adanya dua trait, yaitu sifat Introversi dan sifat Neurotisme, yang masing-masing diungkapkan oleh dua macam metode, pertama metode jawaban ya dan tidak dan kedua metode pilihan ganda. Dengan demikian, dalam contoh ini, terdapat empat macam tes. Hasil pengenaan keempat macam tes tersebut pada sekelompok subjek yang sama kemudian dikorelasikan satu sama lain dan koefisien-koefisien

korelasinya dimasukkan kedalam suatu matriks validasi. Dalam istilah validitas terdapat skala-skala yang memperlihatkan adanya validitas konvergen (*convergent validity*) dan validitas diskriminan (*diskriminant validity*).

c. Validitas Faktorial

Analisis faktorial merupakan kumpulan prosedur metematik yang kompleks guna menganalisis saling hubungan diantara variabel-variabel dan menjelaskan saling hubungan tersebut dalam bentuk kelompok variabel yang terbatas yang disebut faktor. Oleh karena itu, validitas yang didasarkan melalui prosedur analisis faktor disebut validitas faktorial (*faktorial validity*). Pelaksanaan pengujian validitas konstrak melalui prosedur analisis faktor memerlukan pemahaman yang mendalam mengenai komputasi statistika. Kompleksitas komputasi tersebut menyebabkan perlu uraian mengenai konsep analisis faktorial.

Analisis faktor merupakan salah satu dari analisis ketergantungan (*interpedensi*) antar variabel dan merupakan bagian analisis multivariat. Prinsip dasar analisis faktor adalah mengekstraksi sejumlah faktor (*common factor*) dari gugusan variabel awal.

Analisis faktor bertujuan untuk mendapatkan sejumlah kecil faktor komponen utama yang memiliki sifat mampu menerangkan semaksimal mungkin keragaman data. Analisis faktor menggambarkan hubungan kovarians dari beberapa variabel dalam sejumlah kecil faktor. Variabel-variabel ini dapat dikelompokkan menjadi beberapa faktor, dimana variabel-variabel dalam satu faktor mempunyai korelasi yang tinggi sedangkan korelasi dengan

variabel-variabel pada faktor lain relatif kecil. Faktor-faktor tersebut saling independent dan tiap-tiap faktor dapat diinterpretasikan.

Langkah-langkah dalam analisis faktor adalah sebagai berikut : pertama adalah menghitung korelasi antara semua variabel dalam analisis. Karena salah satu tujuan dari analisis faktor adalah untuk memperoleh faktor yang dapat menjelaskan korelasi tersebut, maka variabel harus berkorelasi satu sama lain. Jika korelasi antar variabel adalah kecil maka kemungkinan besar variabel-variabel tersebut terletak dalam variabel yang berbeda. Analisis faktor layak digunakan jika setengah lebih korelasi antar variabel lebih besar dari 0,3 pada nilai absolutnya. Dengan menggunakan *Bartlett test* dapat diuji hipotesis bahwa matrik korelasi adalah matrik identitas. Uji ini membutuhkan asumsi bahwa matrik data berasal dari distribusi normal multivariat.  $H_0$  ditolak jika *level significant* kecil, lebih kecil dari 0,05 yang berarti bahwa matriks korelasi bukan matriks identitas. Jika *level significant* lebih besar dari 0,05 yang berarti matriks korelasi adalah matriks identitas, maka penggunaan model analisis faktor perlu dipertimbangkan kembali. Hal lain yang perlu diperhatikan adalah bahwa matriks korelasi bukan matriks singular (determinannya nol). Jika matriks korelasi adalah singular, maka variabel yang mempunyai korelasi yang besar dapat diwakili oleh salah satu variabel saja.

Kaiser-Meyer-Olkin (KMO) mengukur kecukupan sampling, merupakan perbandingan antara koefisien korelasi dari observasi dengan koefisien korelasi parsial. Nilai KMO yang kecil merupakan indikasi

bahwa penggunaan analisis faktor harus dipertimbangkan kembali, karena korelasi antara variabel tidak dapat diterangkan oleh variabel lain. Karakteristik pengukuran bahwa nilai KMO sebesar 0,90 adalah sangat bagus, 0,80 bagus, 0,70 cukup, 0,60 kurang, 0,50 jelek dan dibawah 0,50 tidak dapat diterima.

Langkah kedua adalah menganalisis faktor *extraction* dan menentukan banyak faktor yang akan digunakan. Komponen utama pertama menerangkan proporsi keragaman data terbesar. Komponen utama kedua menerangkan proporsi keragaman terbesar kedua setelah komponen utama pertama, dan tidak berkorelasi dengan yang pertama.

Langkah ketiga adalah Rotasi faktor, untuk memudahkan interpretasi faktor. Meskipun faktor matriks yang diperoleh menerangkan hubungan antara faktor dan masing-masing variabel, biasanya sulit untuk mengidentifikasi arti dari faktor berdasarkan matriks ini. Sering kali suatu variabel nampaknya tidak mempunyai korelasi dengan faktor manapun. Sedangkan tujuan dari analisis faktor adalah untuk memperoleh faktor yang berisi variabel-variabel yang mempunyai korelasi yang tinggi satu sama lain, maka rotasi analisis faktor bertujuan untuk merotasi matrik *loading* sehingga mudah diinterpretasikan.

Metode rotasi yang digunakan adalah metode rotasi tegak lurus *varimaks* sehingga menghasilkan matriks *loading* baru. Metode ini digunakan jika model faktor mengasumsikan bahwa faktor kesamaanya (*common faktor*) bersifat *independent*. Dengan merotasi matrik *loading* maka setiap variabel

asal akan mempunyai korelasi yang tinggi dengan faktor tertentu saja dan dengan tidak faktor lainnya, sehingga faktor-faktor tersebut saling *independent*. Dengan demikian setiap faktor akan lebih mudah diinterpretasikan.

Di dalam analisis faktor terdapat dua bahasan, yaitu analisis faktor eksplanatori dan analisis faktor konfirmatori. Analisis faktor eksplanatori bersifat eksploratif, umumnya bertujuan untuk mengetahui faktor-faktor penyusun dari sebuah dimensi kehidupan, beberapa faktor yang akan terbentuk belum dapat dilakukan sebelum dianalisis. Sedangkan analisis faktor konfirmatori bersifat apriori berdasarkan landasan teori dan konsep yang dimiliki, dan faktor-faktor yang terbentuk serta variabel laten yang masuk dalam faktor sudah dapat diketahui.

Dalam prosedur analisis faktor, tes yang dipengaruhi oleh faktor-faktor tertentu disebut sebagai tes dipengaruhi oleh faktor-faktor tertentu disebut sebagai tes yang memiliki muatan faktor (*factor loading*) yang tinggi. Muatan faktor merupakan indeks yang arti dan besarnya mirip dengan koefisien korelasi. Bila faktor-faktor tidak berkorelasi satu sama lain maka muatan faktor bukanlah koefisien korelasi. Bila faktor-faktor tidak berkorelasi akan tetapi seringkali diinterpretasikan seakan-akan koefisien korelasi.

Prosedur validasi faktorial yang lebih lengkap menghendaki disertakannya satu rangkaian tes yang telah terbukti berfungsi dalam mengukur faktor yang bersangkutan. Rangkaian tes ini disebut sebagai *marker test* yang seakan-akan menjadi kriteria bagi ada tidaknya validitas tes yang akan diuji. Tes yang

akan diuji akan dikatakan sebagai memiliki validitas faktorial yang baik apabila menunjukkan muatan faktor yang relatif tinggi sebagaimana muatan pada *marker tests*. Adanya validitas faktorial yang baik juga diperlihatkan oleh rendahnya muatan faktor yang tidak diungkap pada *marker tests*.

d. Validitas Berdasarkan Kriteria

Dalam validitas tes berdasarkan kriteria, umumnya tes yang akan diuji validitasnya disebut sebagai prediktor. Statistik yang akan diperlukan dalam pengujian validitas ini adalah koefisien korelasi antara skor sebagai prediktor dan skor suatu kriteria yang relevan.

e. Validitas Prediktif

Apabila skor kriteria validasi merupakan skor yang hendak diprediksi oleh tes dan karenanya baru dapat diperoleh setelah tenggang waktu tertentu setelah tes dikenakan, maka prosedur validasi berdasar kriteria akan menghasilkan suatu statistik yang disebut koefisien validitas prediktif.

f. Validitas Konkruen

Dalam penyusunan dan pengembangan skala-skala psikologi dan berbagai tes mental, pengujian validitas suatu instrumen dalam menjalankan fungsi ukurnya seringkali dapat dilakukan dengan melihat sejauh mana kesesuaian antara hasil ukur instrumen tersebut dengan hasil ukur instrumen lain yang sudah teruji kualitasnya atau dengan ukuran-ukuran yang dianggap dapat menggambarkan aspek yang diukur tersebut secara reliabel. Dalam hal ini, instrumen yang telah teruji validitasnya atau ukuran yang dianggap tepat itu berlaku sebagai kriteria validasi.

Untuk menguji kesesuaian hasil ukur tes dengan hasil ukur kriteria, data skor keduanya harus diperoleh dari sekelompok subjek. Perhitungan koefisien korelasi antara skor subyek pada tes yang bersangkutan dan skor mereka pada kriterianya akan menghasilkan koefisien korelasi yang merupakan koefisien validitas tes dan disebut koefisien validitas konkruen.

- g. Validitas dan Konfindensi
- h. Variasi Skor-Murni Prediktor

Untuk menjelaskan efek variasi skor-murni terhadap koefisien validitas, dikemukakan sekali lagi bahwa validasi dapat diestimasi dengan cara melihat kesesuaian antara distribusi skor-tampak tes dengan distribusi skor-tampak kriterianya. Dalam berbagai kasus, koefisien yang dihitung dengan cara seperti itu sering kali dipengaruhi oleh suatu keadaan yang disebut restriksi sebaran (*restriction of range*), yaitu keadaan menyempitnya variasi skor dikarenakan berkurangnya jumlah subjek atau dikarenakan sangat homogenya skor subjek. Restriksi sebaran dapat terjadi baik distribusi skor tes sebagai prediktor maupun pada distribusi skor kriteria.

Dalam hal ini apabila  $X$  melambangkan skor tes,  $Y$  melambangkan skor kriteria maka korelasi antara  $X$  dan  $Y$  adalah  $r_{xy}$  yang merupakan koefisien validitas prediktif tes  $X$ . Distribusi skor  $Y$  disebut distribusi marginal sedangkan untuk setiap macam skor  $X$  terdapat distribusi skor  $Y$  masing-masing yang disebut distribusi kondisional. Distribusi skor  $Y$  untuk setiap harga  $X$  mempunyai deviasi standarnya sendiri yang disebut error standar dalam estimasi (*standart error of estimate*) dan diberi simbol  $S_{yx}$ . Hubungan

antara  $r_{xy}$  dan error standar dalam estimasi dapat digunakan untuk menjelaskan efek restriksi sebaran. Error standar dalam estimasi dirumuskan sebagai berikut :

$$s_{yx} = s_y \sqrt{1 - r_{xy}^2}$$

dengan :  $s_y$  = standar distribusi skor  $Y$  ;

$r_{xy}$  = koefisien korelasi skor  $X$  dan skor  $Y$ .

Hubungan antara validitas dan error standar dalam estimasi adalah analog dengan hubungan antara error standar dalam pengukuran (*standard error of measurement*) dengan koefisien reliabilitas  $r_{xx}$ . Apabila terjadi restriksi sebaran dengan berlakunya asumsi *homoscedasticity*, nilai  $S_y^2$  (varians marginal) akan mengecil sedangkan harga  $S_{yx}^2$  tidak dipengaruhi. Oleh karena itu, komponen  $S_{yx}^2 / S_y^2$  akan bertambah besar dan akibatnya tentu saja harga  $r_{xy}$  akan kecil.

Dengan demikian, dapat disimpulkan bahwa adanya restriksi sebaran yang menjadikan variasi skor murni prediktor mengecil akan mengakibatkan *underestimasi* terhadap koefisien validitas yang sesungguhnya. Implikasi kesimpulan ini adalah bahwa pada kelompok subjek yang homogen tidaklah mudah untuk memperoleh koefisien validitas yang tinggi.

i. Koreksi Terhadap Efek Atenuasi

Korelasi antar skor suatu tes dengan skor tes lain tidak dapat lebih tinggi daripada korelasi antara skor tes itu dengan skor murninya sendiri. Dan kuadrat koefisien korelasi skor tes dengan skor murni merupakan koefisien reabilitas ( $r_{xx}' = r_{xt}^2$ ) sedangkan

koefisien korelasi skor tes dengan skor tes lain ( $r_{xy}$ ) merupakan koefisien validitas.

Dengan demikian, tingginya koefisien validitas akan dibatasi oleh tingginya koefisien reliabilitas, yaitu

$$r_{xy} \leq r_{xt} \quad \text{atau} \quad r_{xy}^2 \leq r_{xx'}^2.$$

Koefisien validitas tes  $X$  tidak dapat tinggi apabila reliabilitasnya tidak tinggi. Koefisien validitas tersebut akan lebih rendah lagi apabila kriteria validitasnya juga kurang reliabel. Rendahnya koefisien validitas yang diakibatkan rendahnya reliabilitas tes yang bersangkutan atau oleh rendahnya reliabilitas kriteria validasinya disebut efek atenuasi (*attenuation*). Efek atenuasi menghasilkan *underestimasi* terhadap validitas tes. Estimasi terhadap validitas tes yang sesungguhnya, yaitu apabila efek atenuasi tidak terjadi, dapat dilakukan dengan mengestimasi korelasi antara skor-murni kriterianya. Spearman merumuskan formula koreksi terhadap efek atenuasi sebagai berikut :

$$r_{txty} = \frac{r_{xy}}{\sqrt{r_{xx'}r_{yy'}}$$

dengan :  $r_{xy}$  = koefisien validitas sebelum dikoreksi ;

$r_{xx'}$  = koefisien reliabilitas tes  $X$  ;

$r_{yy'}$  = koefisien reliabilitas kriteria  $Y$ .

Semakin tinggi reliabilitas tes  $X$  dan reliabilitas kriteria validasinya maka koefisien validitas  $X$  akan semakin mendekati estimasi validitas yang sesungguhnya sehingga kita tidak lagi perlu mengadakan koreksi terhadap efek atenuasi.

Koreksi terhadap efek atenuasi dapat juga dilakukan apabila reliabilitas yang rendah hanya terjadi pada salah satu diantara tes atau kriterianya.

j. Validitas dan Panjang Tes

Perubahan banyaknya aitem dalam tes akan menyebabkan perubahan reliabilitas tes itu. Bila aitem dalam tes bertambah banyak, maka sampai batas tertentu reliabilitasnya akan meningkat pula validitasnya. Semakin tinggi proporsi varians skor-tampak yang merupakan varian skor-murni (artinya, semakin tinggi reliabilitas) maka semakin besar pula proporsi varians yang sama-sama dimiliki oleh tes dan kriterianya (artinya, semakin tinggi validitas).

Validitas maksimal suatu tes yang memiliki reliabilitas  $r_{xx'}$  adalah  $\sqrt{r_{xx'}}$ . Bila reliabilitas tes yang telah ditambah jumlahnya aitemnya dilambangkan dengan  $r_{xx'}$  maka validitas maksimalnya diberi lambang  $r_{xy}$ . Perbandingan validitas maksimal antara setelah dan sebelum penambahan aitem adalah :

$$r_{XY(maks)} = \frac{r_{xy(maks)}}{\sqrt{\frac{r_{XX'}}{r_{xx'}}}}$$

Dengan mengetahui validitas dan reliabilitas tes sebelum perubahan jumlah aitem, estimasi terhadap validitasnya setelah perubahan jumlah aitem dapat dihitung dengan formula :

$$r_{XY} = r_{xy} \sqrt{\frac{k}{1 + (k-1)r_{xx'}}}$$

dengan :  $r_{XY}$  = estimasi koefisien validitas setelah perubahan jumlah aitem;

$r_{xy}$  = koefisien validitas sebelum perubahan jumlah aitem;

$r_{xx'}$  = koefisien reliabilitas sebelum perubahan jumlah aitem;

$k$  = perbandingan antara banyaknya aitem setelah dan sebelum perubahan.

Peningkatan validitas yang agak berarti hanya terjadi apabila tes adalah pendek dan koefisien validitasnya belum tinggi. Oleh karena itu, untuk meningkatkan kualitas aitem yang ada daripada sekedar menambah jumlah aitem tanpa disertai usaha untuk memperbaiki aitem-aitem yang kualitasnya masih rendah.

k. Tinggi Koefisien yang Memuaskan

Interpretasi koefisien validitas bersifat relatif. Tidak ada batasan universal yang menunjukkan angka minimal yang harus dipenuhi agar suatu tes dikatakan valid.

Suatu hal yang harus disadari, bahwa dalam estimasi validitas pada umumnya tidak dapat dituntut suatu koefisien yang tinggi sekali sebagaimana halnya dalam interpretasi koefisien reliabilitas. Koefisien validitas yang tidak begitu tinggi, berada disekitar angka 0,5 akan lebih dapat diterima dan dianggap memuaskan daripada koefisien reliabilitas dengan angka yang sama. Namun apabila koefisien validitas itu kurang dari 0,3 biasanya dianggap tidak memuaskan. Angka ini ditetapkan sebagai konvensi yang didasarkan pada asumsi distribusi skor dari kelompok subjek yang berjumlah kurang besar.

Hal yang tidak kurang pentingnya untuk dijadikan pertimbangan adalah sejauhmana tes yang bersangkutan dapat bermanfaat dalam pengambilan keputusan. Seringkali suatu tes yang memiliki koefisien validitas kurang tinggi masih berguna dalam membantu pengambilan keputusan dan dalam kasus-kasus tertentu, misalnya untuk meletakkan subjek pada suatu jenjang relatif dalam kelompoknya,

atau dalam seleksi aitem berdasar indeks validitasnya, atau dalam melihat konsistensi aitem dengan tujuan tes, dan semacamnya.

Terhadap pertanyaan mengenai berapakah koefisien validitas yang dianggap memuaskan, Cronbach mengatakan bahwa jawabannya yang paling masuk akal adalah "yang tertinggi yang dapat kau peroleh". Hal ini dipertegas olehnya dalam kaitan dengan fungsi tes untuk memprediksi hasil suatu prosedur seleksi. Dalam hal ini dikatakan bahwa koefisien yang berkisar antara 0,3 sampai dengan 0,5 telah dapat memberikan kontribusi yang baik terhadap efisien suatu lembaga pelatihan.

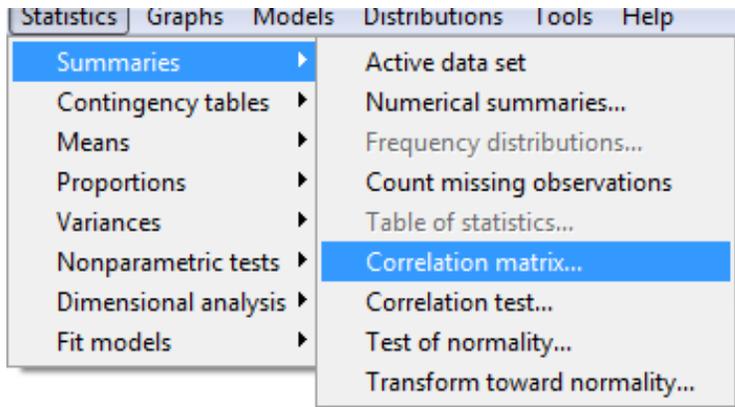
Contoh kasus :

Telah dilakukan survey pendahuluan yang dilakukan kepada 30 responden tentang kepuasan suatu nasabah Bank "XYZ". Adapaun hasil pengumpulan data yang diperoleh adalah :

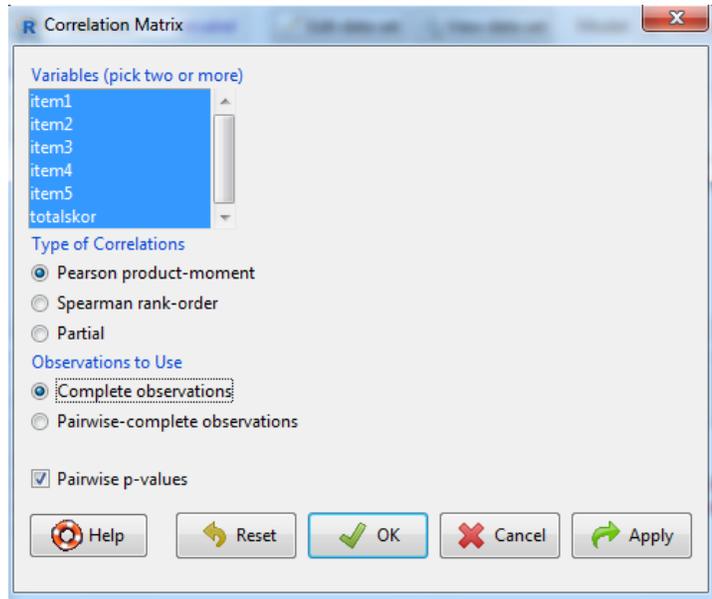
	item1	item2	item3	item4	item5	totalskor
1	7	7	7	6	7	34
2	6	7	7	7	6	33
3	5	5	5	5	4	24
4	6	6	7	6	5	30
5	8	9	9	9	8	43
6	6	6	6	6	7	31
7	8	7	7	7	7	36
8	8	7	8	6	7	36
9	7	6	6	7	5	31
10	7	7	7	8	7	36
11	8	8	7	6	7	36
12	7	7	6	7	6	33
13	7	7	6	5	6	31
14	6	6	6	6	7	31
15	5	6	6	6	5	28
16	8	7	7	7	8	37
17	7	8	7	7	7	36
18	7	7	6	6	5	31
19	6	6	6	6	6	30
20	7	7	7	6	6	33
21	7	7	6	6	6	32
22	8	6	6	6	7	33
23	7	7	6	5	6	31
24	7	7	6	6	7	33
25	7	7	5	5	5	29
26	7	8	6	7	8	36
27	8	9	8	8	8	41
28	6	7	6	7	7	33
29	7	7	8	7	6	35
30	7	6	7	6	6	32

Pertanyaan apakah enam item yang menggambarkan kepuasan nasabah terhadap Bank XYZ tersebut valid ?

Untuk menjawab pertanyaan tersebut, maka dilakukan uji validitas dengan menggunakan rumus product moment pearson antara item 1 sampai dengan item 5 dengan jumlah total skornya. Untuk memudahkan analisis tersebut data diinput kedalam excel kemudian dilakukan import data menggunakan Rcommander dan dilakukan uji validitas dengan bantuan program R paket R-Commander. Dengan langkah-langkah seperti gambar berikut



Selanjutnya akan dianalisis dengan uji korelasi pilih product moment pearson, dari Rcmdr tersebut



Kita akan korelasikan item 1 sampai item 5 dengan total skornya, maka akan muncul dalam output seperti berikut :

```
>rcorr.adjust(latvalrel[,c("item1","item2","item3","item4","item5","totalskor")],
type="pearson", use="complete")
Pearson correlations:
           item1  item2  item3  item4  item5
totalskor
item1      1.0000  0.6320  0.5320  0.3590  0.5981
0.7598
item2      0.6320  1.0000  0.5949  0.5937  0.6463
0.8509
item3      0.5320  0.5949  1.0000  0.6676  0.5259
0.8146
item4      0.3590  0.5937  0.6676  1.0000  0.5777
0.7927
item5      0.5981  0.6463  0.5259  0.5777  1.0000
0.8371
totalskor  0.7598  0.8509  0.8146  0.7927  0.8371
1.0000
```

```

Number of observations: 30
Pairwise two-sided p-values:
      item1  item2  item3  item4  item5
totalskor
item1          0.0002 0.0025 0.0514 0.0005
<.0001
item2    0.0002          0.0005 0.0005 0.0001
<.0001
item3    0.0025 0.0005          <.0001 0.0028
<.0001
item4    0.0514 0.0005 <.0001          0.0008
<.0001
item5          0.0005 0.0001 0.0028 0.0008
<.0001
totalskor <.0001 <.0001 <.0001 <.0001 <.0001
Adjusted p-values (Holm's method)
      item1  item2  item3  item4  item5
totalskor
item1          0.0014 0.0074 0.0514 0.0034
<.0001
item2    0.0014          0.0034 0.0034 0.0010
<.0001
item3    0.0074 0.0034          0.0006 0.0074
<.0001
item4    0.0514 0.0034 0.0006          0.0034
<.0001
item5          0.0034 0.0010 0.0074 0.0034
<.0001
totalskor <.0001 <.0001 <.0001 <.0001 <.0001

```

Dari hasil diketahui bahwa korelasi item 1 sampai dengan item 5 dengan jumlah totalskor menghasilkan nilai sebesar 0.7598, 0.8509 0.8146 0.7927 0.8371 dari hasil perhitungan diperoleh nilai koefisien korelasi, yang selanjutnya dibandingkan dengan r tabel pada n29 dan tingkat kemaknaan 0,05 dan daapt diketahui r tabel sebesar 0.367, sehingga ke lima item tersebut nilai koefisien korelasinya lebih

besar dari ( $>$ ) 0.367. sehingga dapat dikatakan ke lima item tersebut valid semua.

Cara paling praktis dengan melihat nilai  $p$ -value, adalah sebagai berikut :

totalskor  $<.0001$   $<.0001$   $<.0001$   $<.0001$   $<.0001$

pembacaan hasil uji validitas berdasarkan nilai  $p$ -value. Apabila nilai  $p$  value lebih kecil ( $<$ ) dari tingkat kemaknaan 0,05 maka item tersebut valid sebaliknya apabila nilai  $p$ -value  $>$  0,05 maka item alat ukur tersebut tidak valid. Dari harga  $p$ -value contoh uji validitas kuesioner tentang kepuasan nasabah tersebut diketahui bahwa semua nilainya lebih kecil dari 0,05 sehingga kelima item tersebut valid semua.

## 2. Uji reliabilitas

Reliabilitas merupakan penerjemah dari kata *reliability* yang mempunyai asal kata *rely* dan *ability*. Pengukuran yang memiliki reliabilitas tinggi disebut sebagai pengukuran yang reliabel. Nama lain pada reliabilitas antara lain keterpercayaan, keterandalan, keajegan, kestabilan, konsistensi, dan sebagainya, namun pada prinsipnya konsep reliabilitas adalah sejauhmana hasil suatu pengukuran dapat dipercaya.

Pengertian reliabilitas alat ukur dan reliabilitas hasil ukur biasanya dianggap sama, tetapi ada hal yang perlu diperhatikan. Konsep reliabilitas alat ukur erat berkaitan dengan masalah error pengukuran (*error of measurement*) yang menunjukkan pada sejauhmana inkonsistensi hasil pengukuran dilakukan ulang pada kelompok subjek yang sama. Sedangkan konsep reliabilitas hasil ukur erat berkaitan dengan error dalam pengambilan sampel (*sampling error*) yang mengacu kepada inkonsistensi hasil ukur apabila pengukuran dilakukan ulang pada kelompok individu yang berbeda.

Pendekatan konsistensi internal dalam estimasi reliabilitas adalah untuk menghindari masalah-masalah yang biasanya ditimbulkan oleh pendekatan tes ulang dan oleh pendekatan bentuk pararel. Dalam pendekatan konsistensi internal prosedurnya hanya memerlukan satu kali pengenaaan sebuah tes kepada sekelompok individu subjek (*single trial administration*).

Dengan hanya satu kali pengenaaan tes akan diperoleh satu distribusi skor tes dari kelompok subjek yang bersangkutan. Untuk itu, prosedur analisis reliabilitasnya diarahkan pada analisis terhadap aitem-aitem atau terhadap kelompok-kelompok aitem dalam tes itu sehingga perlu dilakukan pembelahan tes menjadi beberapa kelompok aitem yang disebut bagian atau belahan tes. Setiap bagian atau belahan dapat berisi beberapa aitem, bahkan dapat berisi hanya satu aitem saja. Apabila kemudian bagian-bagian tes telah diperoleh maka reliabilitas tes diperlihatkan oleh konsistensi diantara aitem-aitem atau diantara belahan-belahan tes tersebut.

Pembelahan tes dilakukan sedemikian rupa sehingga sedapat mungkin setiap belahan berisi aitem dalam jumlah yang banyak. Akan tetapi, apabila membagi tes kedalam belahan yang berisi aitem dalam jumlah sama banyak tidak mungkin untuk dilakukan, hal itu tidak merupakan masalah lagi dikarenakan sekarang ini telah tersedia rumusan-rumusan baru guna pengujian reliabilitas terhadap tes yang dibelah menjadi bagian-bagian yang berisi aitem dalam jumlah yang tidak seimbang.

Cara pembelahan tes tergantung pula sifat dan fungsi tes jenis skala pengukuran yang digunakan dalam tes bersangkutan. Cara pembelahan itu, pada gilirannya

akan ikut menentukan pula rumusan atau formula mana yang harus digunakan dalam menghitung koefisien reliabilitas.

Suatu tes yang hasilnya sebagian ditentukan oleh kecepatan kerja (*speeded-test*). Misalnya, menghendaki cara pembelahan yang berbeda dari cara pembelahan yang dilakukan terhadap tes yang mengukur kemampuan maksimum (*power-test*). Suatu tes yang berisi aitem-aitem dengan tingkat kesukaran yang sangat bervariasi, tentu tidak setiap karakteristik tes menghendaki cara pembelahan khusus, akan tetapi setiap cara pembelahan tes yang digunakan hendaknya ditekankan pada usaha untuk memperoleh bagian-bagian atau belahan-belahan yang relatif setara.

a. Cara Pembelahan Tes

Membelah suatu tes menjadi beberapa bagian yang setara atau homogen maksudnya adalah mengusahakan agar antara belahan yang satu dengan yang lain memiliki jumlah aitem yang sama banyak, taraf kesukaran yang seimbang, isi yang sebanding, dan sedapat mungkin memenuhi ciri-ciri paralelisme. Walaupun tersedia rumusan guna mengestimasi reliabilitas tes yang belahannya tidak paralel. akan tetapi, estimasi terhadap bagian-bagian yang paralel itu akan lebih meyakinkan bahwa estimasi mendekati nilai reliabilitas yang sesungguhnya, bukan merupakan *underestimasi* (estimasi yang terlalu rendah). Bukan pula *overestimasi* (estimasi yang terlalu tinggi). Berikut adalah beberapa pilihan cara untuk membelah tes menjadi dua bagian, yaitu :

b. Pembelahan Cara Random

Membelah tes menjadi dua bagian secara random dapat dilakukan dengan cara undian sederhana guna menentukan aitem-aitem nomor berapa sajakah yang

dimasukkan menjadi belahan pertama dan mana diikutkan menjadi belahan kedua.

Pembelahan secara random hanya boleh dilakukan apabila tes akan dibelah berisi aitem-aitem yang homogen. Pengertian homogen dalam hal ini harus dipandang dari segi isi (*content homogeneous*) dan juga dari segi taraf kesukarannya apabila tes itu mengukur aspek kognitif. Suatu tes yang berisi aitem-aitem heterogen bila dibelah secara random dapat menghasilkan belahan-belahan yang tidak setara satu sama lain, kecuali apabila tes tersebut terdiri atas aitem yang berjumlah sangat besar.

c. Pembelahan Gasal-Genap

Pembelahan dengan cara gasal-genap (*odd-even splits*) sangat populer dan mudah dilakukan. Dalam cara ini, seluruh aitem yang bernomor urut gasal dijadikan satu kelompok menjadi belahan pertama dan seluruh aitem yang bernomor urut genap dijadikan satu kelompok menjadi belahan kedua. Dengan membelah secara gasal-genap diharapkan akan memperoleh dua bagian yang setara dari segi isi dan taraf kesukaran aitem-aitemnya.

Cara pembelahan ini dapat menghindari kemungkinan terjadinya pengelompokan aitem-aitem tertentu kedalam salah satu belahan saja. Sekalipun, semula aitem-aitem disusun dalam pola urutan tertentu, akan tetapi, sewaktu dilakukan pemisahan gasal-genap maka aitem yang berurutan tadi akan sendirinya terpisah kedalam belahan yang berbeda.

d. Pembelahan Matched-Random Subsets

Untuk tes yang mengukur aspek kemampuan, yang taraf kesukaran aitem serta korelasi aitem dengan skor total tesnya dihitung lebih dahulu, Gulliksen (1950)

mengusulkan suatu cara pembelahan yang disebutnya *match-random subsets*.

Dengan cara ini, setiap aitem dalam tes diletakkan pada satu posisi atau titik tertentu dalam grafik berdasarkan harga indeks kesukara aitem ( $P$ ) dan koefisien korelasi antara aitem yang bersangkutan dengan skor tes ( $r_{ix}$ ). Dengan melihat posisi aitem pada grafik dapat diketahui bahwa setiap aitem yang letaknya berdekatan berarti memiliki karakteristik ( $P$  dan  $r_{ix}$ ) yang relatif sama atau mirip satu sama lain. Kemudian setiap dua aitem yang berdekatan tadi dapat diundi untuk menentukan mana yang dimasukkan kedalam belahan pertama dan mana yang diikutkan kedalam belahan kedua sehingga akan diperoleh dua belahan yang masing-masing berisi sepuluh aitem.

f. Formula Spearman-Brown untuk Belah Dua

Formula Spearman-Brown merupakan sebuah formula komputasi yang sangat populer untuk estimasi reliabilitas tes yang dibelah menjadi dua bagian yang relatif paralel satu dengan yang lain. Formula ini dapat digunakan pada tes yang aitem-aitemnya diberi skor dikotomi maupun bukan dikotomi.

Formula komputasi reliabilitas Sperman-Brown merupakan formula koreksi terhadap koefisien korelasi antara dua bagian tes, dan dirumuskan sebagai berikut :

$$S - B = r_{xx'} = \frac{2(r_{12})}{1 + r_{12}}$$

dimana :  $r_{12}$  = koefisien korelasi antara kedua belahan.

Umumnya untuk memperoleh dua belahan tes yang relatif paralel satusama lain dalam penggunaan formula Spearman-Brown, dilakukan cara pembelahan *gal-genap* atau cara *Matched-Random subsets* dikarenakan

dari dua cara itulah diharapkan akan diperoleh belahan-belahan yang paralel seperti yang dikehendaki.

Skor yang diperoleh subjek dalam tes dihitung terpisah untuk masing-masing belahan sehingga setiap subjek memperoleh dua skor. Kemudian, distribusi skor subjek pada masing-masing belahan dikorelasikan. Koefisien korelasinya dinamai  $r_{12}$ . Estimasi reliabilitas tes diperoleh dengan mengenakan formula Spearman-Brown pada koefisien korelasi antara kedua belahan tersebut.

Formula ini hanya dapat digunakan apabila percaya bahwa asumsi paralelisme diantara kedua belahan terpenuhi. Ciri terpenuhinya asumsi termaksud antara lain adalah apabila kedua belahan tes menghasilkan rata-rata skor (*mean*) yang setara dan varians skor yang sebanding.

Disamping itu, formula Spearman-Brown akan menghasilkan estimasi reliabilitas yang cermat apabila koefisien korelasi diantara kedua belahan tes itu tinggi, karena tingginya korelasi antara kedua belahan merupakan pula indikasi terpenuhinya asumsi paralelisme. Pada kasus yang koefisien korelasi antara keduanya belahan tes tidak begitu tinggi, sebaiknya formula ini tidak digunakan dan beralih kepada cara pendekatan lain.

Metode belahan-dua dapat dipakai untuk mengestimasi kecermatan tes dalam arti ekivalensi (kesetaraan) hasil ukur kedua belahannya. Koefisien ekivalensi ini pada dasarnya sama dengan koefisien reliabilitas.

g. Formula Rulon

Formula Rulon untuk mengestimasi reliabilitas belah-dua tanpa perlu berasumsi bahwa kedua belahan mempunyai varian yang sama.

Perbedaan skor subjek pada kedua belahan tes akan membentuk distribusi perbedaan skor dengan varians yang besarnya ditentukan oleh varians error masing-masing belahan. Karena varians error masing-masing belahan menentukan varians error tes ini dapat diestimasi lewat besarnya varians perbedaan skor diantara kedua belahan. Dengan demikian, dalam melakukan estimasi terhadap reliabilitas tes, varians skor inilah yang perlu diperhitungkan sebagai sumber error. Formula Rulon dirumuskan sebagai berikut :

$$r_{xx'} = 1 - \frac{S_d^2}{S_x^2}$$

dimana :  $S_d^2 =$  varians perbedaan skor kedua belahan;

$S_x^2 =$  varians skor tes.

Formula Rulon juga dapat digunakan pada tes yang aitem-aitemnya diberi skor dikotomi. Koefisien reliabilitas Rulon yang dikenakan pada tes yang telah dibelah menjadi dua bagian ini merupakan estimasi reliabilitas bagi keseluruhan tes sehingga tidak perlu dikenai formula koreksi lagi.

h. Koefisien Alpha ( $\alpha$ ) untuk Belah-Dua

Telah dijelaskan dimuka bahwa formula Spearman-Brown hanya akan menghasilkan estimasi reliabilitas yang cermat apabila belahan-belahan tes yang diperoleh dapat memenuhi asumsi tersebut maka koefisien alpha dapat digunakan.

Walaupun dapat digunakan pada tes yang belahannya tidak paralel satu sama lain, akan tetapi bila kedua belahan tersebut tidak memenuhi asumsi  $\tau$ -*equivalent*, maka koefisien reliabilitas alpha yang diperoleh akan merupakan *underestimasi* terhadap

reliabilitas yang sesungguhnya (artinya, reliabilitas yang sebenarnya mungkin sekali lebih tinggi dari pada koefisien yang diperoleh dari hasil yang diperoleh dari hasil perhitungan). Oleh karena itu, apabila memperoleh hasil perhitungan yang cukup tinggi akan tahu bahwa ada kemungkinan reliabilitas yang sesungguhnya lebih tinggi lagi akan tetapi apabila koefisien yang diperoleh ternyata rendah maka belum dapat dipastikan apakah tes yang bersangkutan memang memiliki reliabilitas rendah ataukah hal tersebut sekedar indikasi tidak terpenuhinya asumsi  $\tau$ -equivalent.

Formula koefisien alpha untuk estimasi reliabilitas tes belah-dua adalah sebagai berikut :

$$r_{xx'} \geq \alpha = 2 \left[ 1 - \frac{S_1^2 + S_2^2}{S_x^2} \right]$$

dengan :  $S_1^2$  dan  $S_2^2$  = varians perbedaan skor kedua belahan;

$$S_x^2 = \text{variens skor tes.}$$

Formula alpha dapat digunakan pada tes yang aitem-aitemnya diberi skor dikotomi.

i. Formula Koefisien Alpha ( $\alpha$ )

Pembelahan tes tidak hanya terbatas pada membagi aitem-aitem tes kedalam dua belahan saja. Cara-cara pembelahan dapat diperluas pemakaiannya untuk membagi tes menjadi beberapa belahan, apabila diperlukan. Bahwa suatu tes yang akan diestimasi reliabilitasnya dapat dibelah menjadi bagian-bagian sebanyak jumlah aitemnya sehingga setiap bagian hanya berisi satu aitem saja. Kecuali dalam kasus tes yang dibelah menjadi belahan sebanyak jumlah aitemnya, hendaknya selalu diupayakan agar belahan-belahan tersebut isinya relatif setara.

Untuk tes yang dibelah menjadi lebih dari dua belahan yang masing-masing berisi aitem dalam jumlah sama banyak dapat menggunakan formula alpha, yang dirumuskan sebagai berikut :

$$\alpha = \left( \frac{k}{k-1} \right) \left[ 1 - \frac{\sum_j S_j^2}{S_x^2} \right]$$

dengan :  $k$  = banyaknya belahan;

$S_j^2$  = varians belahan  $j, j = 1, 2, \dots, k$ .

j. Formula Kuder-Richardson

Apabila suatu tes berisi aitem-aitem yang diberi skor dikotomi sedangkan jumlah aitemnya sendiri tidak begitu banyak, kadang-kadang membagi tes menjadi dua bagian tidak dapat menghasilkan bagian yang setara sedangkan membagi tes menjadi lebih dari dua belahan akan mengakibatkan jumlah aitem dalam setiap belahan terlalu sedikit. Apabila dalam belahan hanya berisi sedikit aitem komputasi reliabilitasnya tidak dapat menghasilkan estimasi yang cermat. Salah satu cara yang dapat dilakukan adalah membelah tes tersebut menjadi sebanyak jumlah aitemnya sehingga setiap belahan berisi hanya satu aitem saja. Estimasi ini dikenal dengan nama formula Kuder-Richardson-20 atau KR-20 atau koefisien  $\alpha$ -20.

Koefisien  $\alpha$ -20 atau KR-20 merupakan rata-rata estimasi reliabilitas dari semua cara belah-dua yang mungkin dilakukan. Koefisien ini juga mencerminkan sejauhmana kesetaraan isi aitem-aitem dalam tes. Formula Kuder-Richardson dirumuskan sebagai berikut :

$$KR - 20 = \left( \frac{k}{k-1} \right) \left[ 1 - \frac{\sum p(p-1)}{S_x^2} \right]$$

dimana :  $k$  = banyaknya aitem dalam tes;

$S_x^2$  = varians skor tes;

$p$  = proporsi subjek yang mendapatkan angka 1 pada suatu aitem, yaitu banyaknya subjek yang mendapat angka 1 dibagi dengan banyaknya seluruh subjek yang melawan aitem.

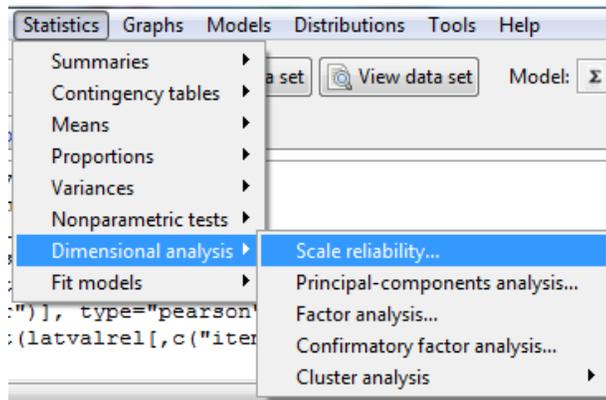
Sedangkan formulasi estimasi reliabilitas yang ke-21, Kuder-Richarson merumuskan sebagai berikut :

$$KR - 21 = \left( \frac{k}{k-1} \right) \left[ 1 - \frac{k\bar{p}(1-\bar{p})}{S_x^2} \right]$$

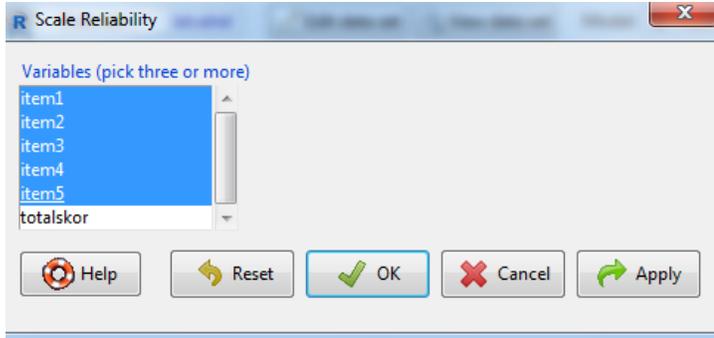
dimana :  $\bar{p}$  = rata-rata  $p$ , yaitu  $\bar{p} = \sum p/k$  .

Untuk contoh kasus diatas data hasil survey kepuasan nasabah sebuah Bank XYZ selanjutnya akan dilakukan uji reliabilitas. Langkah uji reliabilitas dari data uji validitas akan dipilih item alat ukur yang valid saja apabila ada yang tidak valid maka jangan diikutsertakan dalam uji reliabilitas. Sebagai contoh berdasarkan uji validitas diatas ternyata ke lima item tersebut dikatakan valid maka akan kita uji kelima item tersebut reliabilitasnya.

Langkah awal dengan tetap menggunakan data yang sudah tersedia yaitu import data dari excel dengan nama latvalrel kemudian lakukan uji reliabilitas dengan langkah seperti gambar berikut :



Akan tetapi dalam uji reliabilitas ini semua item yang tersedia di uji dan tidak dikut sertakan total skor dari hasil penjumlahan kelima item yang telah dilakukan pada uji validitas, dengan langkah seperti berikut



Maka akan muncul hasil analisis dalam output seperti berikut :

```
>reliability(cov(latvalrel[,c("item1","item2","item3","item4","item5")], use="complete.obs"))
Alpha reliability = 0.8691
Standardized alpha = 0.8702
Reliability deleting each item in turn:
      Alpha Std.Alpha r(item, total)
item1 0.8554      0.8577      0.6344
item2 0.8257      0.8264      0.7598
item3 0.8390      0.8401      0.7036
item4 0.8484      0.8510      0.6651
item5 0.8376      0.8376      0.7148
```

Dari hasil output di atas diperoleh hasil bahwa nilai alfa cronbach sebesar 0,8691 dimana dikatakan reliable jika koefisien yang diperoleh yaitu nilai alfa cronbach lebih besar dari 0.7, jadi kuesioner tentang kepuasan nasabah Bank XYZ tersebut dikatakan reliable.

### C. Rangkuman

Validitas dapat didefinisikan sebagai pendekatan terbaik kebenaran dari proposisi, analisis dan kesimpulan. Banyak orang yang mengatakan tentang validitas riset,

dimana mereka cenderung berfikir pada komponen risetnya., misalnya valid pada alat ukurnya, valid pada sampelnya, valid pada desainnya, namun secara teknik pernyataan – pernyataan tersebut salah, karena alat ukur, sampel dan desain riset tidak mempunyai validitas hanya proposisi yang dapat dikatakan valid (Trochim, 1997). Secara teknik pernyataan-pernyataan diatas seharusnya dinyatakan bahwa pengukuran dan sampel yang benar membuat analisis dan kesimpulan menjadi valid.

Semua penelitian sosial melibatkan pengukuran dan observasi, yang berkonsentrasi pada apa yang ingin diukur. Kesimpulan tentang kualitas pengukuran penelitian memainkan peranan penting dalam hal menentukan kualitas metodologi penelitian itu sendiri, sehingga seringkali pengujian akan validitas lebih dititik beratkan pada pengujian alat ukur.

Reliabilitas dari suatu alat ukur tertuju pada konsistensi atau dapat diulangnya lat ukur pada beberapa fenomena, jika suatu alat ukur reliabel maka alat ukur tersebut dapat mengukur hal yang sama lebih dari satu kali atau menggunakan lebih dari satu metode dan hasilnya sama. Bila kita membicarakan tentang reliabilitas yang dibicarakan adalah skore nilai jawaban responden/observasi bukan individu responden.

Tiga mayor aspek yang ada pada reliabilitas adalah, ekuivalensi, stabilitas, dan homogenitas. Ekuivalensi artinya pengukuran reliabilitas didasarkan pada dua fenomena yang sama , stabilitas artinya kondisi yang ada antara pengukuran yang satu dengan yang lain harus sama atau stabil, sedangkan homogenitas tertuju pada bagaimana kesamaan pengukuran pada masing-masing atribut, semuanya berhubungan dengan hubungan internal atau konsistensi internal.

Hubungan antara reliabilitas dan validitas sangat mudah dipahami dimana suatu alat ukur mungkin saja reliabel namun belum tentu valid, sehingga suatu alat ukur pertama kali harus reliabel sebelum valid, jadi reliabilitas adalah perlu namun belum cukup untuk memenuhi kondisi validitas, dengan kata lain suatu pengukuran mungkin konsisten berhubungan dengan fenomenanya namun kurang mengukur apa yang ingin diukur (tidak valid). Namun perlu diingat disini apabila melakukan suatu proyek penelitian penting sekali bahwa suatu alat ukur harus reliabel dan valid juga, jika tidak maka alat ukur kita pada dasarnya tidak berguna dan menurunkan akurasi dari alat ukur itu sendiri.

Uji Validitas adalah suatu uji yang berguna untuk mengukur sejauh mana suatu alat ukur itu mengukur apa yang ingin diukur.

1. Validitas konstruk: suatu uji yang mengukur apakah alat ukur tersebut sudah benar kerangka konsepnya dan mengukur aspek yang sama, caranya dengan mengukur korelasi setiap item jawaban dengan total jawaban dengan korelasi produk moment:

$$r = \frac{N \left( \sum_{i=1}^n XY \right) - \left( \sum_{i=1}^n X \right) \left( \sum_{j=1}^n Y \right)}{\sqrt{\left[ N \sum_{i=1}^n X^2 - \left( \sum_{i=1}^n X \right)^2 \right] \left[ N \sum_{i=1}^n Y^2 - \left( \sum_{i=1}^n Y \right)^2 \right]}}$$

nilai korelasi r ini akan dibandingkan dengan nilai r pada tabel kritis dengan melihat baris N-2 dengan signifikansi  $\alpha=5\%$  dan  $1\%$ , bilai nilai korelasi tersebut kurang dari nilai r pada tabel angka kritis di Lampiran I (Singarimbun & Effendi, 1989), maka item pertanyaan tersebut tidak valid, atau tidak mengukur aspek yang sama dengan item yang lain, kalau nilai r hitungannya negatif berarti pernyataan tersebut bertolak belakang dengan pernyataan lainnya.

2. Validitas Isi suatu alat pengukur ditentukan oleh sejauh mana isi alat pengukur tersebut mewakili semua aspek yang dianggap semua aspek kerangka konsep. Misalkan untuk mengukur status ekonomi keluarga bisa diukur berdasarkan penghasilan ayah, ibu dan anak.
3. Validitas Eksternal adalah validitas yang diperoleh dengan cara mengkorelasikan alat pengukur baru dengan tolok ukur eksternal, yaitu berupa alat ukur yang sudah valid atau untuk mengukur suatu konsep misal status ekonomi keluarga, dilakukan korelasi antara penghasilan keluarga dengan pernyataan lain yaitu pemilikan barang berharga, atau jenis makanan yang dimakan sehari-hari, bila korelasinya tinggi maka ketiga pernyataan tersebut memiliki validitas eksternal.
4. Validitas prediktif, apakah suatu alat ukur dapat memprediksi apa yang akan terjadi di masa datang, misalkan apakah pilihan seseorang terhadap sesuatu produk dapat mengukur sikap pada suatu produk.
5. Validitas Budaya, suatu alat ukur dapat dikatakan valid bila dapat mengukur responden pada budaya tertentu.
6. Validitas rupa yaitu validitas yang mengukur bentuk dan penampilan suatu alat ukur.

Uji Reliabilitas adalah indeks yang menunjukkan sejauh mana suatu alat pengukur dapat dipercaya atau dapat diandalkan/ menunjukkan konsistensi suatu alat pengukur di dalam mengukur gejala yang sama

1. **Teknik pengukuran ulang;** responden diminta mengisi kusioner sebanyak dua kali dengan selang waktu antara 15-30 hari, lalu hasil pengisian pertama dikorelasikan dengan hasil pengisian kedua dengan korelasi product moment

2. **Teknik belah dua:** teknik ini dilakukan bila terdapat banyak item pertanyaan dalam kuisioner sekitar 50-60 item, makin besar jumlah item makin reliabel.

Caranya:

- a. buang item pertanyaan yang tidak valid
- b. item yang valid bagi jadi dua bagian, bisa acak atau ganjil genap.
- c. Skor untuk setiap bagian dijumlahkan
- d. Korelasikan masing-masing skor total perbagian tadi dengan korelasi produk moment
- e. Hitunglah nilai reliabilitasnya dengan cara memasukkan angka korelasi tadi kedalam rumus

$$r_{.tot} = \frac{2(r_{.tt})}{1 + r_{.tt}}$$

dimana:  $r_{.tot}$  = angka reliabilitas keseluruhan item

$r_{.tt}$  = angka korelasi belahan pertama dan belahan kedua

Teknik tersebut bisa dilakukan dengan membandingkannya dengan nilai

$r$  pada tabel angka kritis pada Lampiran I (Singarimbun & Effendi, 1989), sesuai dengan nilai derajat kebebasan sebesar  $N-2$ .

Teknik lainnya adalah dengan menggunakan metode  $\alpha$  Cronbach dimana setelah skor dibagi dua masing-masing dicari nilai standard deviasi kuadratnya demikian juga dengan standard deviasi kuadrat dari total skor (tanpa dibelah), dan dimasukkan dalam rumus (Cronbach, 1951)

$$\alpha_c = \frac{k}{k-1} \left( 1 - \frac{\sum_{p=1}^k S_p^2}{S_{xt}^2} \right)$$

dimana:

$k$  = jumlah belahan yang dibuat

$S^2_p$  = Standard deviasi skor pada masing-masing belahan

$S^2_{xt}$  = Standard deviasi kuadrat dari total skor

Nilai  $\alpha_c$  adalah nilai  $\alpha$  Cronbach, bilai nilainya diatas 0,7 berarti alat ukur tersebut sudah cukup reliabel.

### 3. Teknik Bentuk pararel (*alternative form*)

Membuat dua buah jenis alat ukur yang mengukur aspek yang sama dan diberikan pada responden yang sama kemudian masing-masing dicari validitasnya, reliabilitasnya dihitung dengan mengkorelasikan antara skor total alat ukur pertama dengan skor total alat ukur kedua.

## D. Tugas

Lakukan pengujian Reliabilitas dan Validitas dari hasil survey pendahuluan yang sudah dilakukan pada 30 orang responden, kemudian berikan analisis dari hasil pengujian anda.

## E. Rujukan

Allen, M.J and Yen, W.M., (1997), *Introduction to Measurement Theory*, Menterey: Books/ Cole Publishing Company.

Azwar,S., (2004), Reliabilitas dan Validitas, Pustaka Pelajar, Yogyakarta.

Cronbach, L.J, (1951), Coefiicient Alpha and The Internal Structure Of Test, *Psychometrika*, 16, 297-334.

Dewi DAAN., (2018). Modul III: Uji Validitas dan Reliabilitas. Statistika Terapan. Universitas Dipenogoro. 1-14.

## F. Bacaan yang Dianjurkan

Allen, M.J and Yen, W.M., (1997), *Introduction to Measurement Theory*, Menterey: Books/Cole Publishing Company.

- Azwar,S., (2004), Reliabilitas dan Validitas, Pustaka Pelajar, Yogyakarta.
- Cronbach, L.J, (1951), Coefficient Alpha and The Internal Structure Of Test, *Psycometrika*, 16, 297-334.
- Dewi DAAN., (2018). Modul III: Uji Validitas dan Reliabilitas. *Statistika Terapan*. Universitas Dipenogoro. 1-14.

# BAB X

## STATISTIKA NON PARAMETRIK

### A. Pendahuluan

Pada Bab ini akan dibahas hal-hal yang mendasar berkenaan dengan statistika non parametrik dimulai dari konsep statistika non parametrik, Uji Wilcoxon Dua Sampel, Uji Wilcoxon Sampel Berpasangan, Uji Kruskal-Wallis dan Uji Friedman dalam persoalan ekonomi dan bisnis.

#### **Tujuan Pembelajaran :**

setelah mempelajari materi pada bab ini mahasiswa diharapkan memahami prinsip konsep statistika non parametric, uji wilcoxon dua sampel uji wilcoxon sampel berpasangan uji kruskal-wallis uji friedman dalam persoalan ekonomi dan bisnis.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nila  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM) (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Uji Wilcoxon Dua Sampel
2. Uji Wilcoxon Sampel Berpasangan
3. Uji Kruskal-Wallis
4. Uji Friedman

## B. Konsep Statistika Non Parametrik

Analisis yang tidak didasarkan atas asumsi distribusi pada data. Umumnya teknik ini dipakai untuk data dengan ukuran kecil sehingga tidak cukup kuat untuk mengasumsikan distribusi tertentu pada data. Digunakan untuk menganalisis data yang berbentuk nominal dan ordinal. Ada banyak macam metode pengujian dalam statistika non parametrik. Akan tetapi, dalam buku ini akan dibahas metode yang sesuai dengan fasilitas yang terdapat di dalam paket R Commander.

Pada uji-uji pada statistika parametrik diterapkan/dipakai sebagai uji statistik apabila (1) skala data/pengukuran sekurang-kurangnya berskala interval, dan (2) data yang dimiliki terdistribusi normal. Namun, apabila salah satu atau kedua syarat pada uji statistik parametrik ini tidak terpenuhi, maka uji-uji pada statistika parametrik tidak dapat dipergunakan/diterapkan, sehingga diperlukan uji-uji lain, selain uji-uji pada statistika parametrik, yakni uji-uji pada statistika non-parametrik.

Kapan uji-uji pada statistik non-parametrik digunakan?

1. Bila hipotesis yang harus diuji tidak melibatkan suatu parameter populasi.
2. Bila data telah diukur dengan skala yang lebih lemah dibanding yang dipersyaratkan oleh uji-uji pada statistik parametrik yang semestinya digunakan. Sebagai contoh, data mungkin terdiri atas data hitung atau data peringkat, sehingga menghalangi penerapan uji-uji pada statistik parametrik yang semestinya lebih tepat.
3. Bila asumsi-asumsi yang diperlukan agar penggunaan suatu uji-uji pada statistik parametrik, misalnya data yang dimiliki terdistribusi normal, menjadi kabur/tidak valid. Dalam banyak hal, rancangan suatu proyek riset mungkin

menganjurkan penggunaan uji-uji pada statistik parametrik tertentu. Bagaimanapun, pemeriksaan data mungkin mengungkapkan bahwa salah satu atau beberapa asumsi yang mendasari pengujian betul-betul tidak bisa dipenuhi/dipatuhi. Dalam hal ini, uji-uji pada statistik non-parametrik acap kali merupakan pengganti satu-satunya.

4. Bila hasil-hasil riset harus segera disajikan dan perhitungan-perhitungan terpaksa dikerjakan secara manual.

### 1. Uji Wilcoxon Dua Sampel

Metode untuk menentukan apakah dua buah populasi independen mempunyai rata-rata yang sama. Uji ini dapat digunakan pada sembarang distribusi kontinu dan merupakan padanan uji T dua sampel independen.

Perumusan Hipotesis :

$H_0 = \mu_1 = \mu_2$ , rata-rata populasi sampel pertama dan kedua adalah sama

$H_1 = \mu_1 \neq \mu_2$ , rata-rata populasi sampel pertama dan kedua adalah tidak sama

Statistik Uji dan keputusan :

Terdapat beberapa persamaan/ rumusan untuk mencari statistik uji yaitu :

#### 1. Walsh Average, caranya :

- Kurangkan nilai hipotesis median dengan nilai masing-masing observation  $Y_i$ .
- Kurangi jumlah observasi dengan jumlah observasi yang sama dengan nilai median.
- Hitung  $n(n + 1) / 2$  buah (Walsh) averages yaitu  $(Y_i + Y_j) / 2$  for  $i \leq j$ .

2. Estimasi Median, caranya :

Urutkan  $n(n + 1) / 2$  buah hasil walsh average, apabila  $n(n + 1) / 2$  ganjil maka penaksir median adalah nilai tengah data, jika tidak maka penaksir median adalah rata-rata 2 data tengah.

Statistik Uji (Wilcoxon Test) : jumlah Walsh Average yang melebihi nilai hipotesis median kemudian ditambahkan 1,5 kali Walsh Average yang nilainya sama dengan nilai hipotesis median.

Atau digunakan pendekatan distribusi Normal :

$$Z_w = \frac{\left| W - \frac{n(n + 1)}{24} \right| - 0,5}{\sqrt{\frac{n(n + 1)(2n + 1)}{24}}}$$

Nilai P-value dihitung dengan pendekatan normal:

Hipotesis Alternatif	P-value
H <sub>1</sub> : Median > Nilai Hipotesis	P (Z <sub>w</sub> > Z)
H <sub>1</sub> : Median < Nilai Hipotesis	P (Z <sub>w</sub> < - Z)
H <sub>1</sub> : Median ≠ Nilai Hipotesis	2 P ( Z <sub>w</sub>   > Z)

Nilai Z didapatkan dari table normal berdasarkan nilai  $\alpha$ .

Keputusan : Apabila nilai P-value lebih besar dari nilai  $\alpha$  maka hipotesis awal (H<sub>0</sub>) diterima begitu pula sebaliknya bila nilai P-value lebih kecil dari  $\alpha$  maka hipotesis awal (H<sub>0</sub>) ditolak. Nilai  $\alpha$  ditentukan sendiri oleh peneliti.

Contoh kasus :

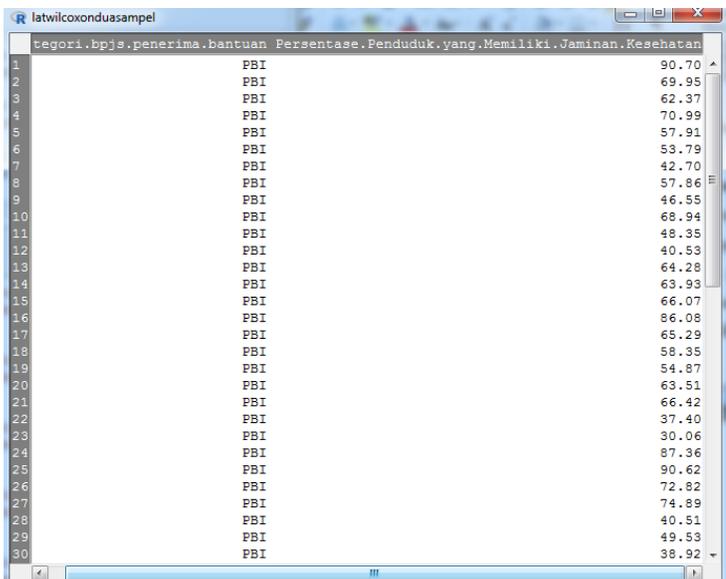
Seorang peneliti ingin melakukan pengujian apakah terdapat perbedaan signifikan antara BPJS Penerima Bantuan Iuran dan BPJS Non-Penerima disuatu provinsi. Gunakan  $\alpha = 5\%$

$H_0$  : ada perbedaan antara BPJS Kesehatan Penerima Bantuan Iuran dan BPJS Non-Penerima Bantuan Iuran.

$H_1$  : tidak ada perbedaan antara BPJS Kesehatan Penerima Bantuan Iuran dan BPJS

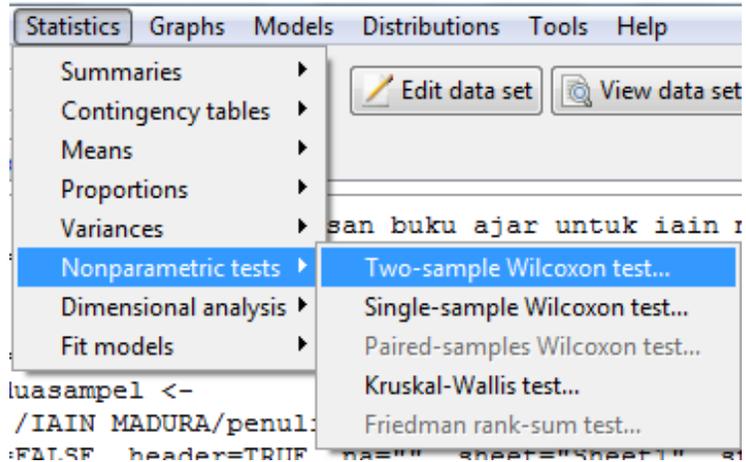
Non-Penerima Bantuan Iuran

Prosedur melakukan pengujian **Uji Wilcoxon Dua Sampel** menggunakan R-Commander adalah menginput data di Excel kemudian melakukan impor data menggunakan R-Commander dengan klik menu Data  $\rightarrow$  impor data  $\rightarrow$  from excel file...misalkan dengan memberi nama data yang akan kita gunakan pada dataset tersebut adalah latwilcoxonduasampel, pilih sheet yang digunakan dalam menginput data di excel. Untuk mengecek data yang kita input dapat dilihat di view data set. Adapun datanya adalah sebagai berikut :

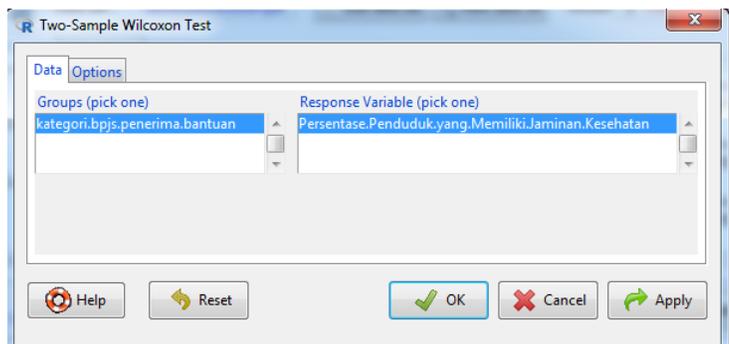


	PBI	Persentase.Penduduk.yang.Memiliki.Jaminan.Kesehatan
1	PBI	90.70
2	PBI	69.95
3	PBI	62.37
4	PBI	70.99
5	PBI	57.91
6	PBI	53.79
7	PBI	42.70
8	PBI	57.86
9	PBI	46.55
10	PBI	68.94
11	PBI	48.35
12	PBI	40.53
13	PBI	64.28
14	PBI	63.93
15	PBI	66.07
16	PBI	86.08
17	PBI	65.29
18	PBI	58.35
19	PBI	54.87
20	PBI	63.51
21	PBI	66.42
22	PBI	37.40
23	PBI	30.06
24	PBI	87.36
25	PBI	90.62
26	PBI	72.82
27	PBI	74.89
28	PBI	40.51
29	PBI	49.53
30	PBI	38.92

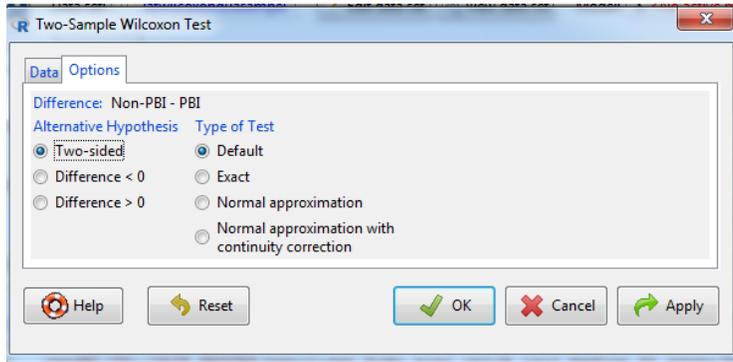
Setelah data sudah siap langkah selanjutnya yaitu melakukan analisis uji sampel berpasangan dengan langkah klik statistics → Nonparametric test → Two Sample Wilcoxon test



Selanjutnya akan terbuka menu Two Sample Wilcoxon test seperti berikut :



Klik option untuk memilih hipotesis alternatifnya dan memilih tipe dari pengujian kemudian klik OK



Sehingga diperoleh output hasil pengujiannya perbedaan rata-rata sampel berpasangan yang diperoleh dari contoh kasus diatas adalah sebagai berikut :

```
>
Tapply(Persentase.Penduduk.yang.Memiliki.Jaminan.Kesehatan ~ kategori.bpjs.penerima.bantuan, median,
na.action=na.omit, data=latwilcoxonduasampel)
+ # medians by group
Non-PBI      PBI
  29.11     62.37
```

```
>
wilcox.test(Persentase.Penduduk.yang.Memiliki.Jaminan.Kesehatan ~ kategori.bpjs.penerima.bantuan,
alternative="two.sided",
data=latwilcoxonduasampel)
```

Wilcoxon rank sum exact test

```
data:
Persentase.Penduduk.yang.Memiliki.Jaminan.Kesehatan by kategori.bpjs.penerima.bantuan
W = 53, p-value = 5.903e-13
alternative hypothesis: true location shift is not equal to 0
```

Dari output R di atas diperoleh nilai statistic  $W = 53$  dengan  $p\text{-value} = 5,903.10^{-13}$ . Dengan demikian  $p\text{-value}$  lebih kecil dari nilai  $\alpha = 0,05$ , sehingga dapat diambil Kesimpulannya  $H_0$  diterima, artinya terdapat perbedaan signifikan antara persentase jumlah peserta BPJS Penerima Bantuan Iuran dan BPJS NonPenerima Bantuan Iuran.

## 2. Uji Wilcoxon Sampel Berpasangan

Teknik ini merupakan penyempurnaan dari uji tanda (*sign test*). Kalau dalam uji tanda besarnya selisih nilai angka antara positif dan negative tidak diperhitungkan sedangkan dalam uji Wilcoxon ini diperhitungkan. Seperti dalam uji tanda teknik ini digunakan untuk menguji signifikansi hipotesis komparatif dua sampel yang berkorelasi apabila datanya berbentuk ordinal (berjenjang).

Apabila sampel pasangan lebih besar dari 25, maka distribusinya akan mendekati distribusi normal. Untuk itu digunakan rumus  $z$  dalam pengujiannya.

$$z = \frac{T - \mu_T}{\sigma_T}$$
 dimana  $T$  adalah jumlah jenjang / ranking yang kecil

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

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Kesimpulan untuk kriteria pengujian hipotesis yaitu  $H_0$  diterima jika harga jumlah jenjang yang terkecil  $T$  dari perhitungan lebih besar dari harga  $T$  tabel atau  $p\text{-value} > \alpha$ .

Contoh kasus :

Suatu penelitian dilakukan untuk melihat apakah terdapat perbedaan yang signifikan pada kinerja

keuangan sebelum dan sesudah go public dengan skala penliatn 1,2,3,4 dan 5 (Tidak Baik, Kurang Baik, Cukup Baik, Baik, dan Sangat Baik). Dengan menggunakan tingkat kemaknaan  $\alpha = 0,05$  .

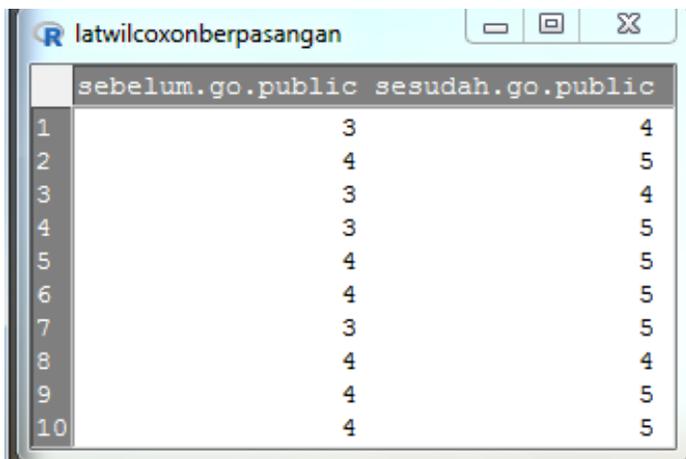
Hipotesis penelitian tersebut adalah :

$H_0$  = tidak ada perbedaan yang signifikan pada kinerja keuangan sebelum dan sesudah go public

$H_1$  = terdapat perbedaan yang signifikan pada kinerja keuangan sebelum dan sesudah go public

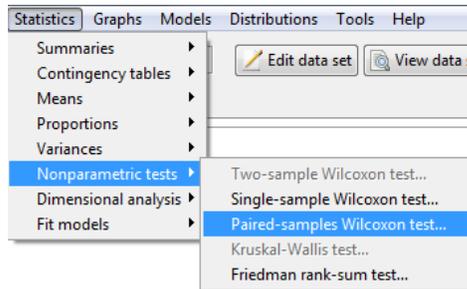
Prosedur melakukan pengujian rata-rata dua sampel saling bebas menggunakan R-Commnander adalah sebagai berikut :

Input data kedalam Excell kemudian melakukan impor data menggunakan R-Commander dengan menu klik Data → impor data → from excel file...untuk melihat data yang kita input sesuai apa tidak dengan klik view data set missal data kita namai didataset latwilcoxonberpasangan. Adapun datanya adalah sebagai berikut :

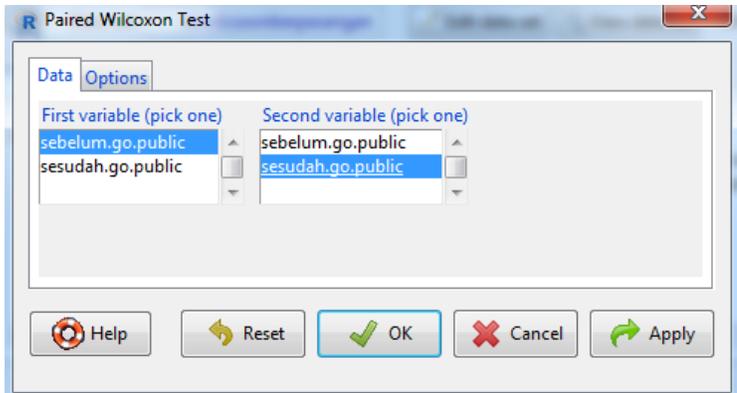


	sebelum.go.public	sesudah.go.public
1	3	4
2	4	5
3	3	4
4	3	5
5	4	5
6	4	5
7	3	5
8	4	4
9	4	5
10	4	5

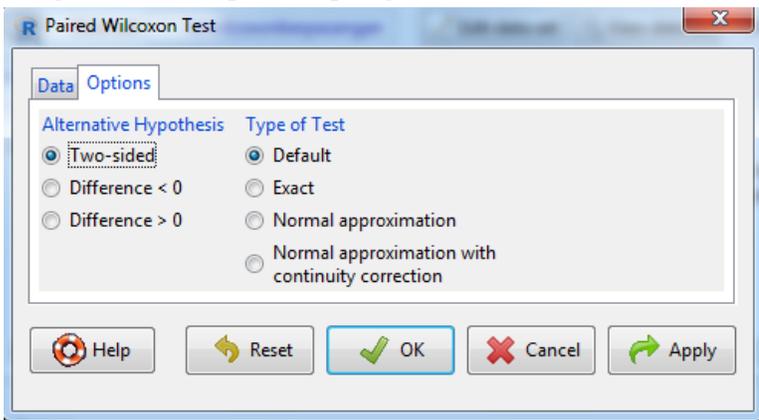
Langkah selanjutnya yaitu melakukan uji wilcoxon sampel berpasangan dengan klik menu statistics → Nonparametric test → Paired Sample Wilcoxon test...



Pilih sebelum go public pada variabel pertama dan sesudah go public pada variabel ke dua



Klik Option untuk menyetting hipotesis alternative yang diinginkan dan tipe dari pengujian, kemudian klik OK



Ouput hasil perhitungan menggunakan R-Commander pengujian Wilcoxon berpasangan untuk kasus ini diperoleh sebagai berikut :

```
> latwilcoxonberpasangan <-  
+   readXL("D:/IAIN MADURA/penulisan buku ajar  
untuk iain madura dr azwar/buku ajar satistika  
bisnis azwar untuk iain madura/latihan wilcoxon  
berpasangan.xls"),  
+   rownames=FALSE, header=TRUE, na="",  
sheet="Sheet1", stringsAsFactors=TRUE)  
> with(latwilcoxonberpasangan,  
median(sebelum.go.public - sesudah.go.public,  
na.rm=TRUE)) # median difference  
[1] -1  
  
> with(latwilcoxonberpasangan,  
wilcox.test(sebelum.go.public, sesudah.go.public,  
alternative='two.sided', paired=TRUE))
```

Wilcoxon signed rank test with continuity  
correction

```
data: sebelum.go.public and sesudah.go.public  
V = 0, p-value = 0.006008  
alternative hypothesis: true location shift is  
not equal to 0
```

Dari output hasil perhitungan Program R tersebut diperoleh nilai p-value sebesar 0,006008. Untuk pengambilan keputusan langsung saja dengan mudah cukup melihat nilai dari p-value dibandingkan dengan nilai tingkat kemaknaan  $\alpha=0,05$ . Sehingga dapat diperoleh bahwa nilai p-value lebih kecil dari nilai  $\alpha=0,05$ , sehingga keputusannya  $H_0$  pengujia tersebut ditolak. Jadi kesimpulannya terdapat perbedaan yang signifikan pada kinerja keuangan sebelum dan sesudah go public.

### 3. Uji Kruskal-Wallis

Merupakan Alternatif uji nonparametrik dari analisis varian satu arah (One Way ANOVA) yang digunakan untuk menguji signifikansi perbedaan mean antara tiga atau lebih sampel independen.

Uji ini merupakan alternatif bagi uji F untuk pengujian kesamaan beberapa nilai mean dalam analisis varian bila ingin menghindari dari asumsi kenormalan data (ordinal). Bila data berbentuk interval atau rasio maka perlu diubah dalam bentuk ordinal terlebih dahulu. Perumusan hipotesis:

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

$$H_1 : \text{minimal ada satu } \mu_i \neq \mu_j ; i \neq j$$

Statistik uji

Dari literatur diketahui ada dua rumus statistik uji yaitu:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

Statistik uji H digunakan pada program paket SPSS.

$$KW = \frac{12}{N(N+1)} \sum_{i=1}^k n_i \left( \bar{R}_i - \frac{N+1}{2} \right)^2$$

$$KW_{ties} = \frac{KW}{1 - \frac{\sum_{i=1}^k (t_i^3 - t_i)}{N^3 - N}}$$

Dimana

$n_i$  = banyaknya nilai pengamatan (ulangan) pada tiap-tiap perlakuan.

$k$  = banyaknya perlakuan yang diuji.

$R_i$  = jumlah ranking tiap perlakuan

$N$  = total pengamatan.

$t$  = frekuensi data yang sama / kembar.

H dan KW akan mendekati distribusi  $\chi^2$  dengan derajat bebas  $k-1$ .

Kesimpulan:

Jika nilai  $H < \chi_{k-1}^2$  atau  $KW < \chi_{k-1}^2$  maka terima hipotesis nol dengan kata lain, semua perlakuan mempunyai nilai mean yang sama.

Jika nilai  $H > \chi_{k-1}^2$  atau  $KW > \chi_{k-1}^2$  maka tolak hipotesis nol dan terima hipotesis alternative dengan kata lain, minimal ada satu perlakuan yang berbeda meannya.

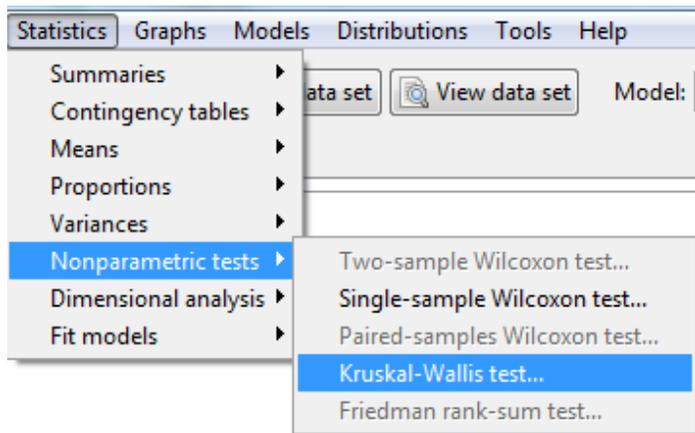
Atau tolak hipotesis nol jika Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} < \alpha$ .

Contoh kasus :

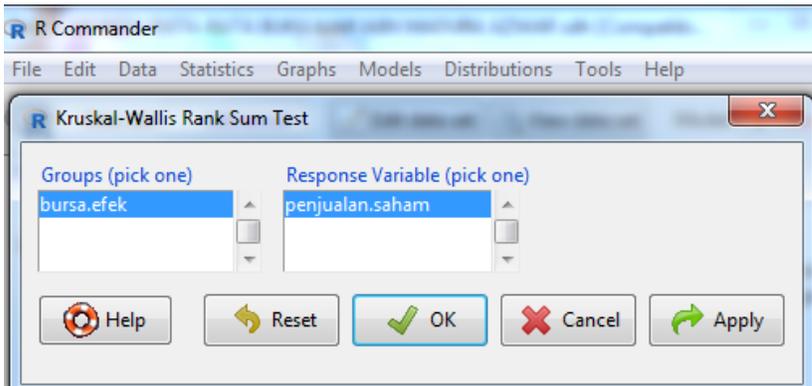
Seorang investor saham ingin mengetahui apakah penjualan saham di tiga bursa efek yang berada di amerika yaitu bursa efek A, bursa efek B dan bursa efek C terdapat perbedaan dalam penjualan saham yang dilakukan. Data penjualan ketiga bursa efek tersebut langsung diinput ke dalam program R dengan menu dengan klik menu Data  $\rightarrow$  impor data  $\rightarrow$  from excel file...misalkan dengan memberi nama data yang akan kita gunakan tersebut adalah latkruswal, pilih sheet yang digunakan dalam menginput data di excel. Untuk mengecek data yang kita input dapat dilihat di view data set. Adapun datanya adalah sebagai berikut :

Row	bursa efek	penjualan.saham
1	bursa efek A	110
2	bursa efek A	103
3	bursa efek A	97
4	bursa efek A	96
5	bursa efek A	105
6	bursa efek A	110
7	bursa efek A	100
8	bursa efek A	93
9	bursa efek A	90
10	bursa efek A	103
11	bursa efek A	97
12	bursa efek A	103
13	bursa efek A	90
14	bursa efek A	97
15	bursa efek A	99
16	bursa efek A	106
17	bursa efek A	90
18	bursa efek A	96
19	bursa efek A	109
20	bursa efek A	90
21	bursa efek B	93
22	bursa efek B	95
23	bursa efek B	110
24	bursa efek B	102
25	bursa efek B	110
26	bursa efek B	91
27	bursa efek B	96
28	bursa efek B	90
29	bursa efek B	105
30	bursa efek B	93

Setelah data sudah siap langkah selanjutnya yaitu melakukan analisis uji kruskal-wallis dengan langkah klik statistics → Nonparametric test → kruskal-wallis test...



Selanjutnya akan terbuka menu uji kruskal-wallis, pilih variabel bursa efek sebagai kelompok dan variabel penjualan saham sebagai variabel respon kemudian klik OK, seperti berikut :



Sehingga diperoleh output hasil pengujian uji kruskal-wallis dari contoh kasus diatas adalah sebagai berikut :

```
> Tapply(penjualan.saham ~ bursa.efek, median,
na.action=na.omit,
+ data=latkruswal) # medians by group
bursa efek A bursa efek B bursa efek C
           98.0           96.0           101.5

> kruskal.test(penjualan.saham ~ bursa.efek,
data=latkruswal)
```

Kruskal-Wallis rank sum test

```
data: penjualan.saham by bursa.efek
Kruskal-Wallis chi-squared = 1.6191, df = 2, p-
value = 0.4451
```

Dari output diatas dapat dilakukan pengujian hipotesis dengan membentuk terlebih dahulu hipotesisnya yaitu:

$H_0$  = Tidak ada perbedaan penjualan saham antara bursa efek A, bursa efek B dan bursa efek C

$H_1$  = ada perbedaan penjualan saham antara bursa efek A, bursa efek B dan bursa efek C

Hasil output menunjukkan bahwa di bursa efek A nilai median dari penjualan traktor sebesar 98 lembar. Untuk penjualan di bursa efek B nilai median penjualan traktor 96 lembar. Sedangkan penjualan di bursa efek C nilai median 101,5 lembar. Nilai Kruskal-Wallis chi-squared yang didapat sebesar 1,6191 dan probabilitas (p) atau p value sebesar 0.445. Kriteria pengujian dengan membandingkan nilai Kruskal-Wallis chi-squared statistik dengan chi-square tabel. Jika nilai chi-square statistik (Kruskal-Wallis chi-squared) < chi-square tabel maka terima  $H_0$  (Hipotesis null). Nilai chi-square tabel pada taraf signifikansi 5% dan df (degree of freedom) 2 sebesar 5,991. Karena nilai chi-square statistik (Kruskal-Wallis chi-squared) 1,6191 < 5,991 chi-square tabel maka terima hipotesis nul ( $H_0$ ). Hal ini didukung juga dengan nilai (p) yang di peroleh sebesar 0.4451 > 0.05. Kesimpulan: Tidak ada perbedaan penjualan saham antara bursa efek A, B dan C.

#### 4. Uji Friedman

Menguji hipotesis komparatif k sampel yang berpasangan bila datanya berbentuk ordinal atau ranking. Bila data yang terkumpul berbentuk interval, atau ratio, maka data tersebut diubah ke dalam bentuk data ordinal. Uji Friedman merupakan alternatif dari ANOVA satu jalur. Uji ini dilakukan jika asumsi-asumsi dalam statistik parametris tidak terpenuhi contohnya asumsi data tidak memenuhi distribusi normal, atau juga karena sampel yang terlalu sedikit.

Menguji keidentikan populasi dalam satu blok. atau Metode untuk mengetahui apakah varians dari perlakuan pada Rancangan Blok Acak homogen. Tes ini menggunakan ANOVA dua arah.

Hipotesis =  $H_0$ : populasi – populasi dalam suatu blok identik.

$H_1$ : paling sedikit ada satu perlakuan yang cenderung menghasilkan nilai-nilai yang lebih besar dibandingkan paling sedikit salah satu perlakuan lain.

Statistik Uji = - Friedman's test tanpa pengulangan :

$$FM = \frac{12b}{k(k+b)} \sum_{i=1}^k \left( Ri - \frac{k+1}{2} \right)^2, \text{ db} = k - 1$$

atau

$$\chi^2 = \frac{12b}{bk(k+1)} \left( \sum_{j=1}^k R_j^2 \right) - 3b(k+1)$$

- Friedman's test dengan pengulangan :

$$FM_{ties} = \frac{b^2}{\sum_{j=1}^b S_{bj}^2} \sum_{i=1}^k \left( Ri - \frac{k+1}{2} \right)^2, \text{ db} = k - 1$$

atau

$$\chi^2_{ties} = \frac{\frac{12b}{bk(k+1)} \left( \sum_{j=1}^k R_j^2 \right) - 3b(k+1)}{1 - \frac{\sum_{i=1}^b T_i}{bk(k^2 - 1)}}$$

dimana :

b = banyak blok.

k = banyak treatment.

$R_j$  = jumlah rank tiap treatment.

$S_{bj}^2$  = varian rank tiap blok.

Kesimpulan =  $x^2_{hit} > x^2_{\alpha(k-1)}$  maka tolak  $H_0$

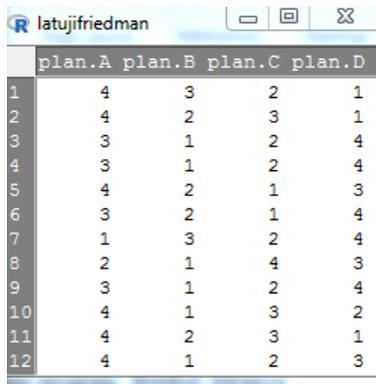
$x^2_{hit} < x^2_{\alpha(k-1)}$  maka terima  $H_0$

Atau tolak hipotesis nol jika Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} < \alpha$ .

Contoh kasus :

Manajer HRD membuat 4 strategi perencanaan perubahan jadwal masuk karyawan. Sebanyak 12 karyawan diambil sebagai sampel dan diminta untuk meranking persetujuan terhadap program. Berikut Prosedur melakukan uji Friedman menggunakan R-Commnander adalah:

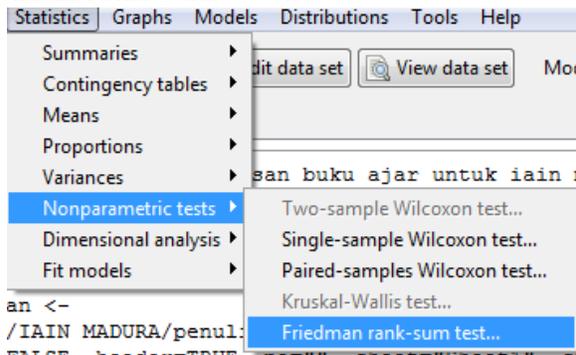
Input data kedalam Excell kemudian melakukan impor data menggunakan R-Commander dengan menu klik Data → impor data → from excel file...untuk melihat



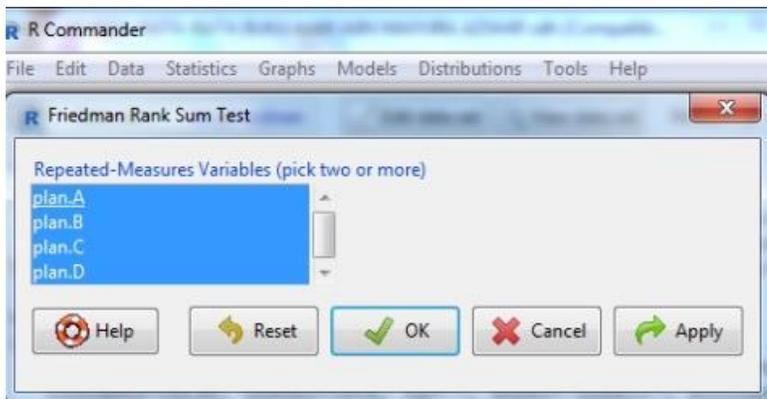
	plan.A	plan.B	plan.C	plan.D
1	4	3	2	1
2	4	2	3	1
3	3	1	2	4
4	3	1	2	4
5	4	2	1	3
6	3	2	1	4
7	1	3	2	4
8	2	1	4	3
9	3	1	2	4
10	4	1	3	2
11	4	2	3	1
12	4	1	2	3

Data yang kita input sesuai apa tidak dengan klik view data set misal data kita namai latujifriedman. Adapun datanya adalah sebagai berikut :

Langkah selanjutnya yaitu melakukan uji friedman dengan klik menu statistics → Nonparametric test → friedman rank-sum test...



Pilih variabel yang akan diuji pada repeated-Measured Variables , yaitu Plan A Plan B Plan C dan Plan D. Kemudian klik OK.



Ouput hasil perhitungan menggunakan R-Commander untuk kasus uji friedman diatas diperoleh sebagai berikut:

```
> local({
+           .Responses      <-
+   na.omit(with(latujifriedman,
+   cbind(plan.A,      plan.B,      plan.C,
+   plan.D)))
+   cat("\nMedians:\n")
+   print(apply(.Responses, 2, median))
+   friedman.test(.Responses)
+ })
Medians:
plan.A plan.B plan.C plan.D
  3.5   1.5   2.0   3.0
  Friedman rank sum test
data:  .Responses
Friedman chi-squared = 10.3, df = 3, p-
value = 0.01618
```

Hipotesis penelitian tersebut adalah

$H_0$  = tidak terdapat perbedaan preferensi yang signifikan terhadap 4 strategi perencanaan perubahan jadwal masuk karyawan.

$H_1$  = terdapat perbedaan preferensi yang signifikan terhadap 4 strategi perencanaan perubahan jadwal masuk karyawan.

Dari output R diatas diperoleh p-value sebesar 0,01618. Dengan demikian p-value lebih kecil dari nilai  $\alpha = 0,05$ , sehingga dapat diambil keputusan bahwa  $H_0$  uji Friedman tersebut ditolak, hal ini berarti terdapat perbedaan preferensi yang signifikan terhadap 4 strategi perencanaan perubahan jadwal masuk karyawan. Karyawan lebih menginginkan rencana A untuk diimplementasikan.

### C. Rangkuman

Analisis yang tidak didasarkan atas asumsi distribusi pada data. Umumnya teknik ini dipakai untuk data dengan ukuran kecil sehingga tidak cukup kuat untuk mengasumsikan distribusi tertentu pada data. Digunakan untuk menganalisis data yang berbentuk nominal dan ordinal.

Pada uji-uji pada statistika parametrik diterapkan/dipakai sebagai uji statistik apabila (1) skala data/pengukuran sekurang-kurang berskala interval, dan (2) data yang dimiliki terdistribusi normal. Namun, apabila salah satu atau kedua syarat pada uji statistik parametrik ini tidak terpenuhi, maka uji-uji pada statistika parametrik tidak dapat dipergunakan/diterapkan, sehingga diperlukan uji-uji lain, selain uji-uji pada statistika parametrik, yakni uji-uji pada statistika non-parametrik.

Kapan uji-uji pada statistik non-parametrik digunakan?

1. Bila hipotesis yang harus diuji tidak melibatkan suatu parameter populasi.
2. Bila data telah diukur dengan skala yang lebih lemah dibanding yang dipersyaratkan oleh uji-uji pada statistik

parametrik yang semestinya digunakan. Sebagai contoh, data mungkin terdiri atas data hitung atau data peringkat, sehingga menghalangi penerapan uji-uji pada statistik parametrik yang semestinya lebih tepat.

3. Bila asumsi-asumsi yang diperlukan agar penggunaan suatu uji-uji pada statistik parametrik, misalnya data yang dimiliki terdistribusi normal, menjadi kabur/tidak valid. Dalam banyak hal, rancangan suatu proyek riset mungkin menganjurkan penggunaan uji-uji pada statistik parametrik tertentu. Bagaimana-pun, pemeriksaan data mungkin mengungkapkan bahwa salah satu atau beberapa asumsi yang mendasari pengujian betul-betul tidak bias dipenuhi/dipatuhi. Dalam hal ini, uji-uji pada statistik non-parametrik acap kali merupakan pengganti satu-satunya.
4. Bila hasil-hasil riset harus segera disajikan dan perhitungan-perhitungan terpaksa dikerjakan secara manual.

Ada banyak macam metode pengujian dalam statistika non parametrik. Akan tetapi, dalam buku ini akan dibahas metode yang sesuai dengan fasilitas yang tersedia di dalam paket R Commander. Uji Wilcoxon dua sampel, metode untuk menentukan apakah dua buah populasi independen mempunyai rata-rata yang sama. Uji ini dapat digunakan pada sembarang distribusi kontinu dan merupakan padanan uji T dua sampel independen. Uji Wilcoxon Sampel Berpasangan Teknik ini merupakan penyempurnaan dari uji tanda (*sign test*). Kalau dalam uji tanda besarnya selisih nilai angka antara positif dan negative tidak diperhitungkan sedangkan dalam uji Wilcoxon ini diperhitungkan. Seperti dalam uji tanda teknik ini digunakan untuk menguji signifikansi hipotesis komparatif dua sampel yang berkorelasi apabila datanya berbentuk ordinal (berjenjang).

Uji Kruskal-Wallis Merupakan Alternatif uji nonparametrik dari analisis varian satu arah (One Way ANOVA) yang digunakan untuk menguji signifikansi perbedaan mean antara tiga atau lebih sampel independen. Uji ini merupakan alternatif bagi uji F untuk pengujian kesamaan beberapa nilai mean dalam analisis varian bila ingin menghindar dari asumsi kenormalan data (ordinal). Bila data berbentuk interval atau rasio maka perlu diubah dalam bentuk ordinal terlebih dahulu. Uji Friedman Menguji hipotesis komparatif k sampel yang berpasangan bila datanya berbentuk ordinal atau ranking. Bila data yang terkumpul berbentuk interval, atau ratio, maka data tersebut diubah ke dalam bentuk data ordinal. Uji Friedman merupakan alternative dari ANOVA satu jalur. Uji ini dilakukan jika asumsi-asumsi dalam statistik parametris tidak terpenuhi contohnya asumsi data tidak memenuhi distribusi normal, atau juga karena sampel yang terlalu sedikit.

Menguji keidentikan populasi dalam satu blok. atau Metode untuk mengetahui apakah varians dari perlakuan pada Rancangan Blok Acak homogen. Tes ini menggunakan ANOVA dua arah.

#### **D. Tugas**

Lakukan pengujian Uji Wilcoxon Dua Sampel, Uji Wilcoxon Sampel Berpasangan, Uji Kruskal-Wallis dan Uji Friedman dari hasil pencarian data yang sudah dilakukan pada, kemudian berikan analisis dari hasil pengujian anda.

#### **E. Rujukan**

- Conover, W.J. (1999). *Practical NonParametric stistics Third Edition*. John Wiley & Sons, Inc.
- Daniel, Wayne W. (1989). *Statistika Non Parametrik Terapan*. Jakarta: Gramedia.

- Djarwanto. (2003). Statistik Nonparametrik. Bpfe-Yogyakarta
- Praptono. (1986). Metode Statistika Nonparametrik. Depdikbud UT.
- Siegel, Sidney. (1985). Statistik Nonparametrik untuk Ilmu-ilmu Sosial. (diterjemahkan oleh Peter Hagul). Jakarta: Gramedia.
- Sugiyono. (2001). Statistik Nonparametrik untuk Penelitian. Cv Alfabeta Bandung.

#### **F. Bacaan yang Dianjurkan**

- Conover, W.J. (1999). Practical NonParametric statistics Third Edition. John Wiley & Sons, Inc.
- Daniel, Wayne W. (1989). Statistika Non Parametrik Terapan. Jakarta: Gramedia.
- Djarwanto. (2003). Statistik Nonparametrik. Bpfe-Yogyakarta
- Praptono. (1986). Metode Statistika Nonparametrik. Depdikbud UT.
- Siegel, Sidney. (1985). Statistik Nonparametrik untuk Ilmu-ilmu Sosial. (diterjemahkan oleh Peter Hagul). Jakarta: Gramedia.
- Sugiyono. (2001). Statistik Nonparametrik untuk Penelitian. Cv Alfabeta Bandung.

# BAB XI

## ANALISIS KORELASI

### A. Pendahuluan

Pada Bab ini akan dibahas tentang analisis korelasi yang terdiri dari metode korelasi product moment pearson, korelasi spearman rank dan korelasi kendall tau beserta prakteknya dalam persoalan ekonomi dan bisnis.

#### **Tujuan pembelajaran :**

Setelah mempelajari materi pada bab ini mahasiswa diharapkan mampu memahami dan menerapkan analisis korelasi yang terdiri dari metode korelasi product moment pearson, korelasi spearman rank dan korelasi kendall tau beserta prakteknya dalam persoalan ekonomi dan bisnis..

#### **Capaian Pembelajaran :**

1. Sikap dan tata nila  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM) ). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

- 1) Korelasi product moment pearson,
- 2) Korelasi spearman rank
- 3) Korelasi kendall tau

## B. Konsep Analisis Korelasi

Korelasi dapat diartikan sebagai hubungan. analisis korelasi bertujuan untuk mengetahui pola dan keeratan hubungan antara dua atau lebih variabel. Sehingga dapat disimpulkan konsep analisis korelasi yaitu Menguji hubungan antara dua variabel atau lebih yang ada pada sampel untuk diberlakukan pada seluruh populasi dimana sampel diambil.

Tujuan dalam analisis korelasi yaitu Menguji hipotesis apakah terdapat hubungan yang signifikan atau tidak antara dua variabel atau lebih.

Dengan bentuk hipotesis

$H_0$  : tidak terdapat hubungan yang signifikan

$H_1$  : terdapat hubungan yang signifikan

Arah hubungan antara dua variabel dapat dibedakan menjadi:

- a. *Direct correlation (positive correlation)*. Perubahan pada satu variabel diikuti perubahan variabel yang lain secara teratur dengan arah gerakan yang sama. Positif bila nilai suatu variabel ditingkatkan maka akan meningkatkan variabel yang lain dan sebaliknya bila diturunkan akan menurunkan nilai variabel yang lain
- b. *Inverse correlation (negative correlation)*. Perubahan pada satu variabel diikuti perubahan variabel yang lain secara teratur dengan arah gerakan yang berlawanan. Hubungan yang negatif, bila nilai suatu variabel dinaikkan maka akan menurunkan nilai variabel lain dan sebaliknya.
- c. *Nihil correlation*. Arah hubungan kedua variabel yang tidak teratur.

Persoalan akan timbul jika kita berhadapan dengan pertanyaan apakah ada hubungan antara variabel-variabel dari sekumpulan data yang sedang kita selidiki. Penyelidikan untuk mengetahui hubungan antara dua variabel biasanya diawali dengan usaha untuk menemukan bentuk terdekat dari

hubungan tersebut dengan cara menyajikannya dalam bentuk diagram pencar (scatter plot). Diagram ini menggambarkan titik-titik pada bidang X dan Y dimana setiap titik ditentukan oleh pasangan nilai X dan Y.

Koefisien korelasi sering dilambangkan dengan huruf (r). Mencari seberapa besar hubungan antara dua variabel atau lebih. Angka yang menunjukkan arah dan kuatnya hubungan antara dua variabel atau lebih. Koefisien korelasi dinyatakan dengan bilangan, bergerak antara 0 sampai +1 atau 0 sampai -1. Apabila korelasi mendekati +1 atau -1 berarti terdapat hubungan yang kuat, sebaliknya korelasi yang mendekati nilai 0 bernilai lemah. Apabila korelasi sama dengan 0, antara kedua variabel tidak terdapat hubungan sama sekali. Pada korelasi +1 atau -1 terdapat hubungan yang sempurna antara kedua variabel. Notasi positif (+) atau negatif (-) menunjukkan arah hubungan antara kedua variabel. Pada notasi positif (+), hubungan antara kedua variabel searah, jadi jika satu variabel naik maka variabel yang lain juga naik. Pada notasi negatif (-), kedua variabel berhubungan terbalik, artinya jika satu variabel naik maka variabel yang lain justru turun.

Berdasarkan hasil nilai koefisien korelasi kemudian terdapat nilai Tingkat korelasi yang digunakan untuk menunjukkan seberapa besar kekuatan yang ada pada nilai tersebut, adapun klasifikasinya adalah sebagai berikut :

- ▶ 0,00 - 0,199 = sangat rendah
- ▶ 0,20 - 0,3999 = rendah
- ▶ 0,40 - 0,5999 = sedang
- ▶ 0,60 - 0,799 = kuat
- ▶ 0,80 - 1,000 = sangat kuat

## 1. Korelasi Product Moment Pearson

Menguji signifikansi hipotesis asosiatif bila masing-masing variabel yang dihubungkan berbentuk interval atau rasio dengan sampel memenuhi asumsi distribusi normal (parametrik).

Rumus dari korelasi product moment pearson yaitu :

$$r = \frac{N \left( \sum_{i=1}^n X_i Y_i \right) - \left( \sum_{i=1}^n X_i \right) \left( \sum_{j=1}^n Y_j \right)}{\sqrt{\left[ N \sum_{i=1}^n X_i^2 - \left( \sum_{i=1}^n X_i \right)^2 \right] \left[ N \sum_{i=1}^n Y_i^2 - \left( \sum_{i=1}^n Y_i \right)^2 \right]}}$$

Dengan kriteria pengambilan kesimpulan yaitu

$H_0$  Ditolak jika Nilai tingkat kemaknaan yang diperoleh (p-value)  $< \alpha$  Dengan terlebih dahulu menentukan nilai taraf signifikansinya yaitu  $\alpha$ . Atau bisa menggunakan kriteria pengujian berdasarkan nilai  $r$  dengan kriteria jika  $-r_{\text{tabel}} \leq r_{\text{hitung}} \leq r_{\text{tabel}}$ , maka  $H_0$  diterima

Bandingkan  $r_{\text{hitung}}$  dengan  $r_{\text{tabel}}$  jika  $r_{\text{hitung}} > r_{\text{tabel}}$  maka  $H_0$  ditolak untuk mencari nilai  $r_{\text{tabel}}$  perhatikan derajat kebebasan yaitu  $n-2$  ( $n$  adalah banyaknya sampel).

Contoh kasus

Dilakukan penelitian untuk mengetahui adakah hubungan antara biaya promosi dengan nilai penjualan suatu produk. Gunakan tingkat kemaknaan 0,05.

Penyelesaian :

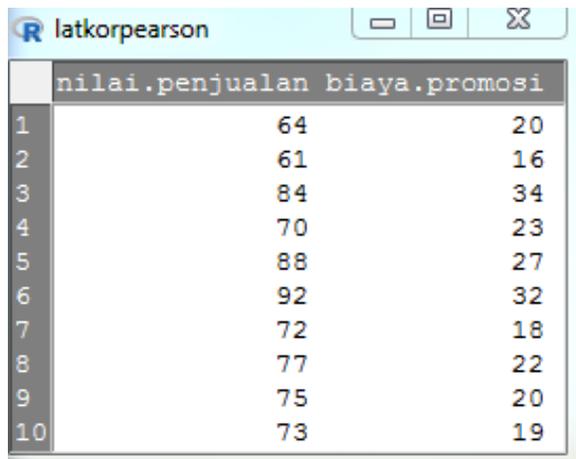
Hipotesis pengujian tersebut dalam bentuk kalimat adalah

$H_0$  : Tidak terdapat hubungan yang positif dan signifikan antara variabel Biaya Promosi dengan Nilai Penjualan.

H<sub>1</sub> : Terdapat hubungan yang positif dan signifikan antara variabel Biaya Promosi dengan Nilai Penjualan.

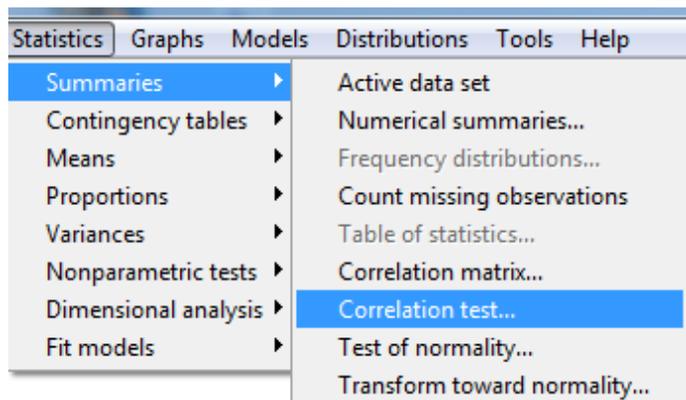
Prosedur perhitungan menggunakan bantuan R-Commander adalah :

Meninput data hasil penelitian ke Excel kemudian lakukan impor data kedalam R-commander dengan klik menu Data → impor data → from excel file...masukkan nama untuk data set misalkan latkorpearson. Hasilnya seperti berikut ini...

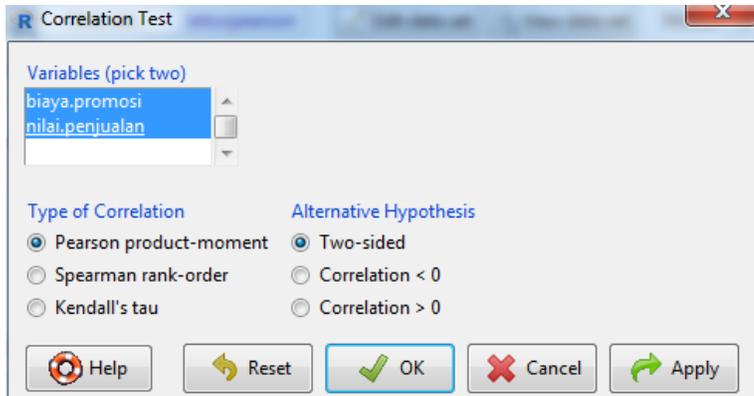


	nilai.penjualan	biaya.promosi
1	64	20
2	61	16
3	84	34
4	70	23
5	88	27
6	92	32
7	72	18
8	77	22
9	75	20
10	73	19

Selanjutnya siap melakukan uji korelasi menggunakan metode product moment pearson dengan menu R-comander seperti berikut :



Pilih variable biaya promosi dan nilai penjualan yang akan diuji hubungannya kemudian pilih type of Correlation menggunakan Pearson Product-Moment kemudian klik OK.



Output yang diperoleh dari kasus diatas adalah sebagai berikut :

```
> with(latkorpearson, cor.test(biaya.promosi,
  nilai.penjualan, alternative="two.sided",
  method="pearson"))
Pearson's product-moment correlation
data: biaya.promosi and nilai.penjualan
t = 4.4183, df = 8, p-value = 0.002232
alternative hypothesis: true correlation is not
equal to 0
95 percent confidence interval:
 0.4525692 0.9618097
sample estimates:
      cor
0.8422086
```

Berdasarkan output tersebut diperoleh nilai korelasi sebesar 0,8422086 sehingga dapat dikatakan berkorelasi secara sangat kuat sedangkan untuk menguji hubungannya dengan melihat nilai p-value yaitu 0,002232 dimana nilai tersebut lebih kecil dari 0,05 maka  $H_0$  harus ditolak sehingga dapat disimpulkan Terdapat hubungan yang positif dan signifikan antara variabel Biaya Promosi dengan Nilai Penjualan.

## 2. Korelasi Spearman Rank

Ada kalanya kita ingin mengukur kuatnya hubungan antara dua variabel tidak berdasarkan pasangan nilai data yang sebenarnya, tetapi berdasarkan rankingnya. Hubungan tersebut dinamakan *rank correlation coefficient*. Analisis korelasi Spearman termasuk dalam statistik non-parametrik. Metode korelasi ini ditemukan Carl Spearman pada tahun 1904. Sehingga dalam hal ini merupakan Korelasi yang Menguji signifikansi hipotesis asosiatif bila masing-masing variabel yang dihubungkan berbentuk ordinal dengan sampel yang digunakan sedikit

Rumus

$$\rho = 1 - \frac{6 \left( \sum_{i=1}^n d_i^2 \right)}{n(n^2 - 1)}$$

dengan keterangan  $r$  adalah nilai koefisien korelasi spearman rank,  $n$  adalah banyaknya pasangan data dan  $d$  adalah selisih dari tiap pasangan ranking.

Dengan kriteria pengujian hipotesis yaitu  $H_0$  diterima apabila nilai  $\rho_{hitung}$  lebih kecil dari  $\rho_{tabel}$ . atau dengan menggunakan Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} > \alpha$

Dengan menggunakan uji signifikansi yang lain dapat

$$z_h = \frac{\rho}{\frac{1}{\sqrt{n-1}}}$$

menggunakan rumus  $z$  yaitu

pengambilan keputusan hipotesis nol ditolak jika nilai  $Z_{hitung}$  lebih besar dari  $Z_{tabel}$ ,

atau menggunakan  $H_0$  Ditolak jika Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} < \alpha$ .

Contoh kasus

Suatu penelitian data diambil di Perusahaan mebel Kreasi Utama dengan tujuan penelitian ingin mengetahui

seberapa kuat hubungan antara harga dengan pendapatan bulanan perusahaan. Dengan menggunakan tingkat signifikansi 0,05.

Penyelesaian :

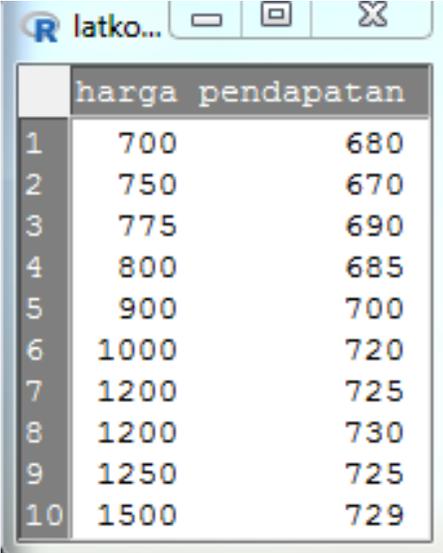
Hipotesis pengujian tersebut dalam bentuk kalimat adalah

$H_0$  : Tidak terdapat hubungan yang positif dan signifikan antara variabel harga dengan pendapatan bulanan perusahaan

$H_1$  : Terdapat hubungan yang positif dan signifikan antara variabel harga dengan pendapatan bulanan perusahaan

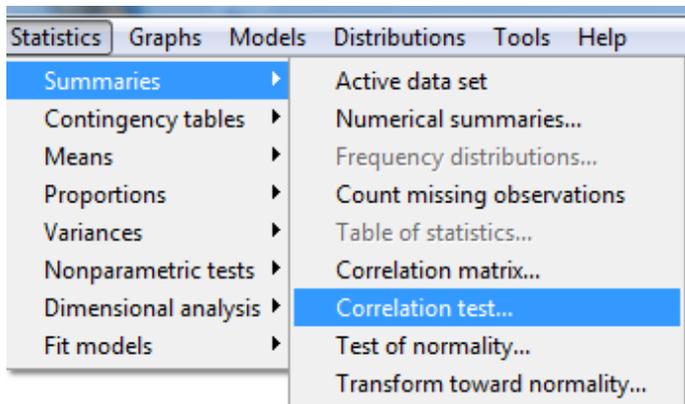
Prosedur perhitungan menggunakan bantuan R-Commander adalah :

Meninput data hasil penelitian ke Excel kemudian lakukan impor data kedalam R-commander dengan klik menu Data → impor data → from excel file...masukkan nama untuk data set misalkan latkorspearman. Hasilnya seperti berikut ini...

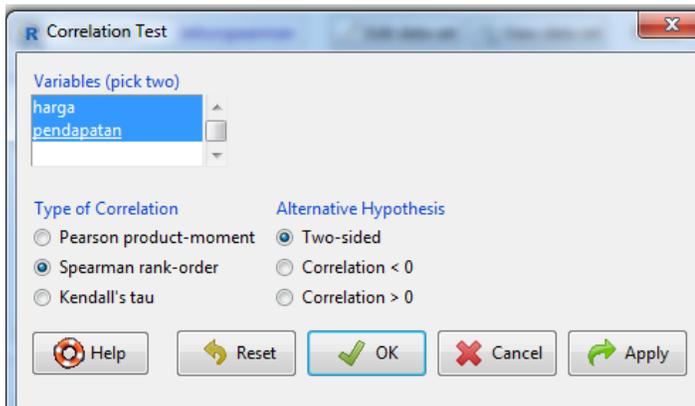


	harga	pendapatan
1	700	680
2	750	670
3	775	690
4	800	685
5	900	700
6	1000	720
7	1200	725
8	1200	730
9	1250	725
10	1500	729

Selanjutnya siap melakukan uji korelasi menggunakan metode spearman rank dengan menu R-commander seperti berikut :



Pilih variable harga dan pendapatan yang akan diuji hubungannya kemudian pilih type of Correlation menggunakan spearman rank kemudian klik OK.



Output yang diperoleh dari kasus diatas adalah sebagai berikut :

```
> with(latkorspearman, cor.test(harga,
  pendapatan, alternative="two.sided",
  method="spearman"))
Spearman's rank correlation rho
data: harga and pendapatan
S = 13.582, p-value = 0.0001817
alternative hypothesis: true rho is not equal to
0
```

sample estimates:

rho

0.9176829

Interpretasinya yaitu dengan melihat Angka koefisien korelasi rho adalah 0,9176892, artinya hubungan antara harga dengan pendapatan sangat erat dan kuat. Koefisien korelasi bertanda positif (+), artinya hubungannya searah sehingga jika harga dinaikkan maka pendapatan perusahaan juga akan mengalami kenaikan. Nilai p-value sebesar 0,0001817 dimana nilainya lebih kecil dari tingkat kemaknaan yaitu 0,05 sehingga menunjukkan bahwa hipotesis nol ditolak dapat disimpulkan Terdapat hubungan yang positif dan signifikan antara variabel harga dengan pendapatan bulanan perusahaan.

### 3. Korelasi Kendall Tau

Selain koefisien korelasi Spearman, terdapat metode analisis korelasi lain yang menguji keeratan hubungan antara variabel X dan Y dimana X dan Y tidak terdistribusi normal atau tidak diketahui distribusinya. Metode ini disebut *Kendall rank correlation*. *Korelasi Kendall Tau* juga didasarkan atas ranking data. Korelasi ini diberi simbol ( $\tau$ ) sehingga dapat digunakan untuk Menguji signifikansi hipotesis asosiatif bila masing-masing variabel yang dihubungkan berbentuk ordinal dengan sampel yang digunakan banyak.

$$\tau = \frac{\sum RA - \sum RB}{\frac{N(N-1)}{2}}$$

Dan dirumuskan sebagai berikut

dengan keterangan  $\sum RA$  adalah jumlah ranking kelompok atas dan  $\sum RB$  adalah jumlah ranking kelompok bawah dan  $N$  adalah banyaknya sampel data yang digunakan.

Karena menggunakan data yang banyak sehingga dapat dikatakan distribusinya mendekati distribusi normal sehingga uji signifikansi koefisien menggunakan rumus Z

$$z = \frac{\tau}{\sqrt{\frac{2(2N + 5)}{9N(N - 1)}}}$$

sebagai berikut :

Dengan kriteria pengujian hipotesis  $H_0$  diterima apabila harga  $Z_{hitung}$  lebih kecil dari tabel dan  $H_1$  diterima apabila harga  $Z_{hitung}$  lebih besar atau sama dengan  $Z_{tabel}$ . atau dengan menggunakan  $H_0$  Ditolak jika Nilai tingkat kemaknaan yang diperoleh yaitu  $p\text{-value} < \alpha$

Contoh kasus :

Seorang peneliti mengadakan sebuah studi untuk merancang dan menguji suatu metodologi untuk menentukan standar-standar untuk kerja penjualan (sales performance) secara analitik. Untuk itu mereka menghimpun data tentang benchmark achievement dan management rating untuk 25 kawasan penjualan. Mereka menghitung benchmark achievement dengan cara membagi sales volume dengan benchmark sales , dan management rating ditentukan berdasarkan motivasi dan usaha yang ditunjukkan oleh masing-masing tenaga penjual.

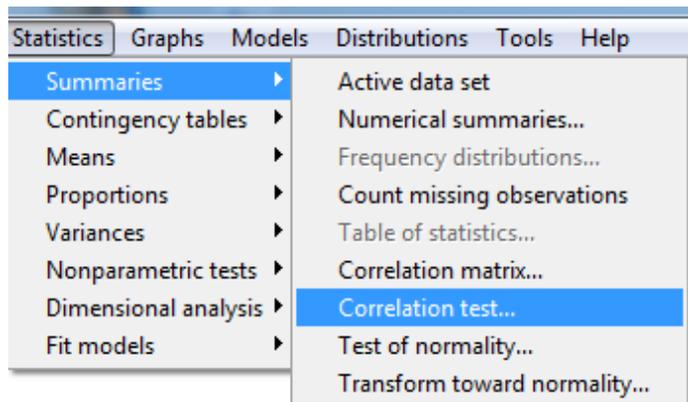
Apakah dari data ini memberikan bukti yang cukup untuk menyimpulkan bahwa benchmark achievement dan management rating memiliki korelasi yang linier (menggunakan taraf signifikansi 0,05)?

Penyelesaian :

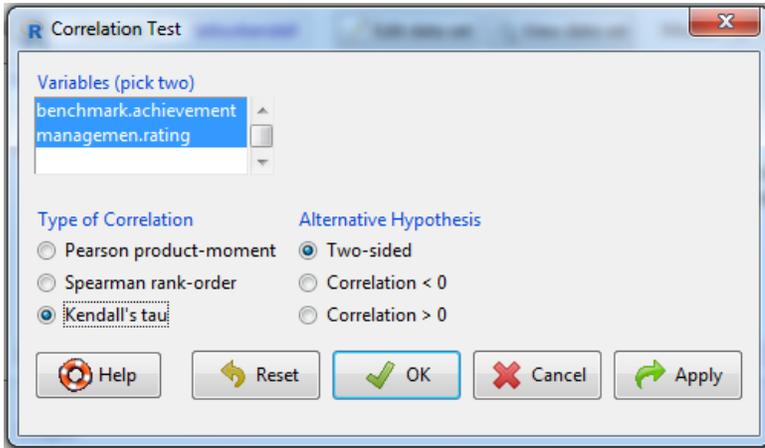
Terlebih dahulu melakukan input data di excell kemudian impor data dengan klik menu Data → impor data → from excel file...

	benchmark.achievement	managemen.rating
1	2	4
2	9	7
3	7	20
4	23	17
5	5	5
6	17	7
7	16	6
8	25	24
9	4	3
10	10	21
11	20	18
12	15	9
13	8	8
14	11	10
15	1	1
16	21	14
17	14	15
18	3	11
19	13	13
20	18	19
21	22	25
22	19	16
23	24	23
24	6	22
25	12	12

Perhitungan Uji Korelasi, menggunakan metode korelasi Kendal Tau dapat dilakukan dengan R-Commander, yaitu menggunakan menu Statistika, pilih Summaries, pilih Correlation test...



Setelah itu akan terlihat tampilan pilihan tentang dua variabel yang akan diuji korelasinya dan pilih Kendall's Tau kemudian klik OK



Output hasil pengujian korelasi diperoleh seperti berikut :

```
> with(latkorkendall,
  cor.test(benchmark.achievement,
  managemen.rating, alternative="two.sided",
  method="kendall"))
Kendall's rank correlation tau
data: benchmark.achievement and managemen.rating
z = 3.2472, p-value = 0.001165
alternative hypothesis: true tau is not equal to
0
sample estimates:
tau
0.4641075
```

Hipotesis uji korelasi tersebut adalah

$H_0$  : Benchmark achievement dan management rating tidak berkorelasi linier

$H_1$  : Benchmark achievement dan management rating berkorelasi linier

Berdasarkan output diatas diperoleh nilai p value sebesar 0,001165 dimana nilainya lebih kecil dari tingkat kemaknaan yang digunakan yaitu 0,05 maka dapat

disimpulkan  $H_0$  ditolak jadi dapat disimpulkan Benchmark achievement dan management rating berkorelasi linier secara positif karena diperoleh nilai tau sebesar 0,4641075 dan dapat dikatakan masuk kategori berkorelasi sedang.

### C. Rangkuman

Korelasi dapat diartikan sebagai hubungan. analisis korelasi bertujuan untuk mengetahui pola dan keeratan hubungan antara dua atau lebih variabel. Sehingga dapat disimpulkan konsep analisis korelasi yaitu Menguji hubungan antara dua variabel atau lebih yang ada pada sampel untuk diberlakukan pada seluruh populasi dimana sampel diambil. korelasi product moment pearson Menguji signifikansi hipotesis asosiatif bila masing-masing variabel yang dihubungkan berbentuk interval atau rasio dengan sampel memenuhi asumsi distribusi normal (parametrik). Ada kalanya kita ingin mengukur kuatnya hubungan antara dua variabel tidak berdasarkan pasangan nilai data yang sebenarnya, tetapi berdasarkan rankingnya. Hubungan tersebut dinamakan *rank correlation coefficient*. Analisis korelasi Spearman termasuk dalam statistik non-parametrik. Metode korelasi ini ditemukan Carl Spearman pada tahun 1904. Sehingga dalam hal ini merupakan Korelasi yang Menguji signifikansi hipotesis asosiatif bila masing-masing variabel yang dihubungkan berbentuk ordinal dengan sampel yang digunakan sedikit. Selain koefisien korelasi Spearman, terdapat metode analisis korelasi lain yang menguji keeratan hubungan antara variabel X dan Y dimana X dan Y tidak terdistribusi normal atau tidak diketahui distribusinya. Metode ini disebut *Kendall rank correlation*. *Korelasi Kendall Tau* juga didasarkan atas ranking data. Korelasi ini diberi simbol ( $\tau$ ) sehingga dapat digunakan untuk Menguji

signifikansi hipotesis asosiatif bila masing-masing variabel yang dihubungkan berbentuk ordinal dengan sampel yang digunakan banyak.

#### **D. Tugas**

Lakukan pengujian analisis korelasi menggunakan metode korelasi product momento Pearson, korelasi spearman Rank dan korelasi kendall tau dari hasil pencarian data yang sudah dilakukan, kemudian berikan analisis dari hasil pengujian anda.

#### **E. Rujukan**

Mason, R.D & Douglas A. Lind. (1996). Teknik Statistik Untuk Bisnis dan Ekonomi. Penerbit Erlangga, Jakarta.

Usman, H. dan R. Purnomo Setiady Akbar. (2000). Pengantar Statistika. Jakarta : Bumi Aksara.

#### **F. Bacaan yang Dianjurkan**

Mason, R.D & Douglas A. Lind. (1996). Teknik Statistik Untuk Bisnis dan Ekonomi. Penerbit Erlangga, Jakarta.

Usman, H. dan R. Purnomo Setiady Akbar. (2000). Pengantar Statistika. Jakarta : Bumi Aksara.

## **BAB XII**

### **ANALISIS REGRESI**

#### **A. Pendahuluan**

Pada Bab ini akan dibahas analisis regresi sederhana dan analisis regresi berganda, serta aplikasinya dalam persoalan ekonomi dan bisnis menggunakan bantuan program R.

#### **Tujuan pembelajaran :**

Setelah mempelajari materi pada bab ini mahasiswa diharapkan menerapkan analisis regresi sederhana dan analisis regresi berganda, bisa menyelesaikan menggunakan program R.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nilai  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM) ). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. analisis regresi sederhana
2. analisis regresi berganda

#### **B. Konsep Analisis Regresi**

Analisis regresi adalah teknik statistika yang berguna untuk memeriksa dan memodelkan hubungan diantara variabel-variabel. Penerapannya dapat dijumpai secara luas di

banyak bidang seperti teknik, ekonomi, manajemen, ilmu-ilmu biologi, ilmu-ilmu sosial, dan ilmu-ilmu pertanian. Pada saat ini, analisis regresi berguna dalam menelaah hubungan dua variabel atau lebih, dan terutama untuk menelusuri pola hubungan yang modelnya belum diketahui dengan sempurna, sehingga dalam penerapannya lebih bersifat eksploratif.

Dalam analisis regresi diasumsikan bahwa  $x_i$  dalam contoh acak  $\{(x_i, y_i); i = 1, 2, \dots, n\}$ . Bersifat tetap dan bukan merupakan nilai peubah acak. Seandainya suatu contoh lain yang berukuran  $n$  diambil dengan menggunakan nilai-nilai  $x$  yang sama, dapat dibayangkan bahwa nilai-nilai  $y$  akan bervariasi, berbeda dari nilai-nilai sebelumnya. Dengan demikian  $y_1$  dalam pasangan  $(x_1, y_1)$  merupakan suatu nilai peubah acak  $Y_i$ . Untuk memudahkan peubah acak  $Y$  padanan nilai tertentu  $x$  akan dilambangkan dengan  $Y|x$ , sedangkan nilai tengah dan ragamnya masing-masing akan dilambangkan dengan  $\mu_{Y|x}$  dan  $\sigma^2_{Y|x}$ . Jadi jelaslah, jika  $x = x_D$  maka lambang  $Y|x_i$  menyatakan peubah acak  $Y_i$  dengan nilai tengah  $\mu_{Y|x_i}$  dan ragam  $\sigma^2_{Y|x_i}$ .

Kurva yang menghubungkan nilai tengah setiap sebaran itu disebut kurva regresi. Bila semua nilai tengah  $\mu_{Y|x}$  jatuh tepat pada suatu garis lurus maka regresinya linear dan dapat dinyatakan dalam persamaan

$$\mu_{Y|x} = \alpha + \beta x$$

Parameter  $\alpha$  dan  $\beta$  disebut koefisien regresi. Dengan melambangkan nilai dugaannya masing-masing dengan  $a$  dan  $b$  maka kita menduga  $\mu_{Y|x}$  dengan  $y$  dari regresi contohnya.

$$\hat{y} = a + bx$$

Untuk sembarang nilai  $x$  tertentu, setiap pengamatan  $(x_i, y_i)$  dalam contoh dipenuhi hubungan

$$y_i = \mu_{Y|X_i} + \varepsilon_i$$

Sedangkan dalam hal ini  $\varepsilon_i$  adalah galat acak yang merupakan simpangan vertical titik tersebut dari garis regresi populasinya. Dari asumsi sebelumnya mengenai  $Y_i$ , maka  $\varepsilon_i$  haruslah merupakan nilai suatu peubah acak yang mempunyai nilai tengah nol dan ragam  $\sigma^2$ . Untuk garis regresi contoh, dapat ditulis

$$y_i = \hat{y}_i + e_i$$

Suatu bagian esensial dalam setiap analisis regresi adalah penyusunan selang kepercayaan dan pengujian hipotesis mengenai koefisien regresi  $\alpha$  dan  $\beta$ . Tetapi pertama-tama, ragam  $\sigma^2$  yang tidak diketahui harus diduga dari data contoh. Suatu nilai dugaan takbias bagi  $\sigma^2$  yang didasarkan pada  $n-2$  derajat bebas diberikan oleh rumus

$$s_e^2 = \frac{JKG}{n-2}$$

Suatu nilai dugaan takbias bagi  $\sigma^2$  dengan  $n-2$  derajat bebas diberikan oleh rumus

$$s_e^2 = \frac{n-1}{n-2} (s_y^2 - b^2 s_x^2)$$

Analisis regresi dikelompokkan dari mulai yang paling sederhana sampai yang paling rumit, tergantung tujuan yang berlandaskan pengetahuan atau teori sementara, bukan asal ditentukan saja. Adapun jenisnya ada empat antara lain :

#### 1. Regresi Linier Sederhana

Regresi linier sederhana bertujuan mempelajari hubungan linier antara dua variabel. Dua variabel ini

dibedakan menjadi variabel bebas ( $X$ ) dan variabel tak bebas ( $Y$ ). Variabel bebas adalah variabel yang bisa dikontrol sedangkan variabel tak bebas adalah variabel yang mencerminkan respon dari variabel bebas.

## 2. Regresi Berganda

Regresi berganda seringkali digunakan untuk mengatasi permasalahan analisis regresi yang melibatkan hubungan dari dua atau lebih variabel bebas. Pada awalnya regresi berganda dikembangkan oleh ahli ekonometri untuk membantu meramalkan akibat dari aktivitas-aktivitas ekonomi pada berbagai segmen ekonomi. Misalnya laporan tentang peramalan masa depan perekonomian di jurnal-jurnal ekonomi (*Business Week, Wal Street Journal*, dll), yang didasarkan pada model-model ekonometrik dengan analisis berganda sebagai alatnya.

## 3. Regresi Kurvilinier

Regresi kurvilinier seringkali digunakan untuk menelaah atau memodelkan hubungan fungsi variabel terikat ( $Y$ ) dan variabel bebas ( $X$ ) yang tidak bersifat linier. Tidak linier bisa diartikan bilamana laju perubahan  $Y$  sebagai akibat perubahan  $X$  tidak konstan untuk nilai-nilai  $X$  tertentu. Kondisi fungsi tidak linier ini (kurvilinier) seringkali dijumpai dalam banyak bidang. Misal pada bidang pertanian, bisa diamati hubungan antara produksi padi dengan taraf pemupukan Phospat. Secara umum produksi padi akan meningkat cepat bila pemberian Phospat ditingkatkan dari taraf rendah ke taraf sedang. Tetapi ketika pemberian dosis Phospat diteruskan hingga taraf tinggi, maka tambahan dosis Phospat tidak lagi diimbangi kenaikan hasil, sebaliknya terjadi penurunan hasil. Dalam hal ini, prosedur analisis regresi kurvilinier merupakan prosedur yang sesuai untuk digunakan.

4. Regresi Dengan Variabel *Dummy* (Boneka)  
Analisis regresi tidak saja digunakan untuk data-data kuantitatif (misal : dosis pupuk), tetapi juga bisa digunakan untuk data kualitatif (misal : musim panen). Jenis data kualitatif tersebut seringkali menunjukkan keberadaan klasifikasi (kategori) tertentu, sering juga dikategorikan variabel bebas ( $X$ ) dengan klasifikasi pengukuran nominal dalam persamaan regresi. Sebagai contoh, bila ingin meregresikan pengaruh kondisi kemasan produk dodol nenas terhadap harga jual. Pada umumnya, cara yang dipakai untuk penyelesaian adalah memberi nilai 1 (satu) kalau kategori yang dimaksud ada dan nilai 0 (nol) kalau kategori yang dimaksud tidak ada (bisa juga sebaliknya, tergantung tujuannya). Dalam kasus kemasan ini, bila kemasannya menarik diberi nilai 1 dan bila tidak menarik diberi nilai 0. Variabel yang mengambil nilai 1 dan 0 disebut variabel *dummy* dan nilai yang diberikan dapat digunakan seperti variabel kuantitatif lainnya.
5. Regresi Logistik (*Logistic Regression*)  
Bila regresi dengan variabel bebas ( $X$ ) berupa variabel *dummy*, maka dikategorikan sebagai regresi *dummy*. Regresi logistik digunakan jika variabel terikatnya ( $Y$ ) berupa variabel masuk katagori klasifikasi. Misalnya, variabel  $Y$  berupa dua respon yakni gagal (dilambangkan dengan nilai 0) dan berhasil (dilambangkan dengan nilai 1). Kondisi demikian juga sering dikategorikan sebagai regresi dengan respon biner. Seperti pada analisis regresi berganda, untuk regresi logistik variabel bebas ( $X$ ) bisa juga terdiri lebih dari satu variabel.

## 1. Analisis Regresi Sederhana

Persamaan matematik yang memungkinkan kita meramalkan nilai-nilai suatu peubah bebas disebut persamaan regresi. Dalam hal ini akan dibicarakan masalah pendugaan atau peramalan nilai peubah takbebas  $Y$  berdasarkan peubah bebas  $X$  yang telah diketahui nilainya. Maka data setiap anggota populasi dapat dinyatakan dalam koordinat  $(x,y)$ . Suatu contoh acak berukuran  $n$  dari populasi dapat dilambangkan dengan  $\{(x,y); i = 1,2,3,\dots, n\}$ . Data tersebut kemudian ditebarkan atau diplotkan sehingga dihasilkan apa sebuah diagram pencar. Dengan mengamati diagram pencar ini, apabila terlihat bahwa titik-titiknya mengikuti suatu garis lurus, menunjukkan bahwa bahwa kedua peubah tersebut saling berhubungan secara linear. Bila hubungan linear demikian ini ada, maka akan diusahakan menyatakan secara matematik dengan sebuah persamaan garis-lurus yang disebut garis regresi linear. Dari aljabar atau ilmu ukur analitik diketahui bahwa sebuah garis lurus dapat dituliskan dalam bentuk

$$\hat{y} = a + bx$$

yang dalam hal ini  $a$  menyatakan intersep atau perpotongan dengan sumbu tegak, dan  $b$  adalah kemiringan atau gradiennya. Lambang  $\hat{y}$  digunakan disini untuk membedakan antara nilai ramalan yang dihasilkan garis regresi dan nilai pengamatan  $y$  yang sesungguhnya untuk nilai  $x$  tertentu.

Sekali diputuskan akan digunakan persamaan regresi linear, maka akan menghadapi masalah bagaimana memperoleh rumus untuk menentukan nilai dugaan titik bagi  $a$  dan  $b$  berdasarkan data contoh. Untuk itu akan digunakan prosedur yang disebut metode kuadrat terkecil. Diantara semua kemungkinan garis

lurus yang dapat dibuat pada diagram pencar, metode kuadrat terkecil memilih suatu garis regresi yang membuat jumlah kuadrat jarak vertikal dari titik-titik pengamatan ke garis regresi tersebut sekecil mungkin.

Jadi, bila  $e_1$  menyatakan simpangan vertical dari titik ke- $i$  ke garis regresi, maka metode kuadrat terkecil menghasilkan rumus untuk menghitung  $a$  dan  $b$  sehingga jumlah kuadrat semua simpangan itu minimum. Jumlah kuadrat semua simpangan ini disebut jumlah kuadrat galat sekitar garis regresi dan dilambangkan dengan JKG. Jadi, bila diberikan segugus data berpasangan  $\{(x_i, y_i); i = 1, 2, \dots, n\}$ , maka harus ditentukan  $a$  dan  $b$  sehingga meminimumkan

$$JKG = \sum_{i=1}^n e_1^2 = \sum_{i=1}^n (y_i - a - bx_i)^2$$

Pendugaan Parameter

Bila diberikan data contoh  $\{(x_i, y_i); i = 1, 2, \dots, n\}$ , maka nilai dugaan kuadrat terkecil bagi parameter dalam garis regresi

$$y = a + bx$$

Dapat diperoleh rumus

$$b = \frac{n \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}$$

Dan

$$a = \bar{y} - b\bar{x}$$

#### a. Pengujian Dalam Regresi Linier

Adapun pengujian dalam regresi linear adalah sebagai berikut:

- i. Uji kecukupan data

Dari hasil output data dapat diperlihatkan hasil dari *P-value* lebih kecil dari nilai  $\alpha = 0,05$ , maka dari data tersebut dapat digunakan suatu analisis regresi sehingga tidak diperlukan adanya penambahan sampel. Apabila *P-value* lebih besar daripada nilai  $\alpha$ . maka dari data tersebut tidak dapat digunakan dalam suatu analisis regresi sehingga diperlukan adanya penambahan sampel. Uji kecukupan data diperlukan sekali sebelum melakukan regresi linier.

ii. Uji koefisien determinasi

Dari data output dapat diperlihatkan nilai dari  $R^2$ , hal ini membuktikan bahwa kekuatan dalam menjelaskan perubahan atau pergerakan variabel x terhadap variabel y. Sedangkan  $R^2(\text{adjusted})$  merupakan koefisien determinasi yang menjelaskan banyaknya variabel bebas dari data tersebut. Nilai  $R^2$  dinyatakan dengan rumus sebagai berikut :

$$R^2 = \frac{JK_{regresi}}{JK_{total}}$$

$$= \frac{b \sum X y - n \bar{X} \bar{y}}{\sum y^2 - n \bar{y}^2}$$

$R^2_{\text{adjusted}}$  dinyatakan dengan rumus sebagai berikut :

$$R^2_{\text{adjusted}} = 1 - (1 - R^2) \left( \frac{n-1}{n-k} \right)$$

b. Uji koefisien regresi

Uji ini juga tidak kalah pentingnya dengan uji -uji yang lain. Dari data di atas dapat digunakan pengujian dua cara yaitu pengujian serentak dengan menggunakan F dan pengujian parsial atau individu dengan t.

- Pengujian secara serentak (dengan F)

Diketahui :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$H_1$  : Paling tidak ada satu  $\beta_j \neq 0, j = 1, 2, \dots, k$

Dalam hal ini digunakan pengujian dengan menggunakan tabel distribusi F. Dimana F :

$$F_{\text{hitung}} = \frac{RK_{\text{Reg}}}{RK_{\text{Res}}}$$
$$= \frac{(b'X'y - n\bar{y}^2) / k}{(y'y - b'X'y) / (n - k - 1)}$$

dengan derajat bebas regresi  $k = 1$  dengan derajat pembilang ( $v_1$ ) dan derajat penyebut ( $v_2$ ) yaitu  $n - k - 1$ . Sedangkan apabila nilai dari  $f$  hitung lebih kecil daripada  $F_{\text{table}} (\alpha; k, n - k - 1)$  maka  $H_0$  ditolak.

- Pengujian individu atau parsial (dengan  $t$ )

Diketahui :

$$H_0 : \beta_j = 0$$

$H_1 : \beta_j \neq 0, j = 1, 2, \dots, k$

Rumus atau statistik yang digunakan adalah yang digunakan adalah :

$$|t_{\text{hit}}| = \frac{b_j}{S(b_j)}$$

Apabila :

- a.  $t_{\text{hitung}} > t_{(\alpha/2, n-k-1)}$  maka  $H_0$  ditolak, artinya ada pengaruh  $X_i$  yang sesuai dengan koefesien regresi terhadap model.
- b.  $t_{\text{hitung}} \leq t_{(\alpha/2, n-k-1)}$  maka gagal tolak  $H_0$ , artinya tidak ada pengaruh  $X_i$  yang sesuai dengan koefesien regresi terhadap model.

Contoh kasus :

Seorang peneliti ingin melakukan penelitian dengan ingin mengetahui apakah terdapat pengaruh yang signifikan antara jumlah produk yang terjual di dua kota selama tahun 2018 dengan biaya promosi yang dikeluarkan.

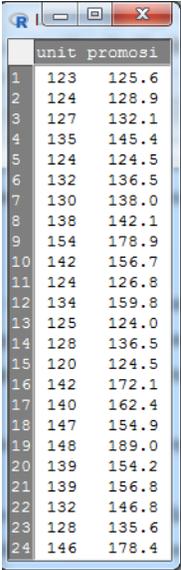
Penyelesaian :

Prosedur langkah-langkah menyelesaikan regresi linier sederhana menggunakan R-Commander yaitu:

Klik menu data → impor data → from excel file...

Masukkan nama untuk dataset (misalkan latregsed) kemudian klik OK.

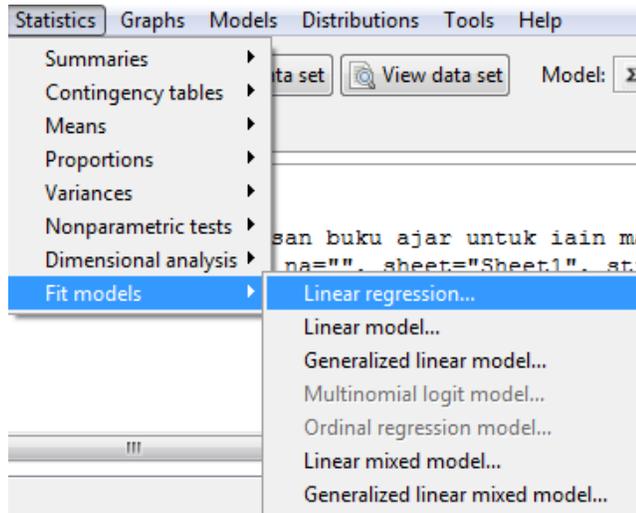
Cek data yang telah di impor dengan melihat di menu view data set. Berikut adalah data yang akan digunakan:



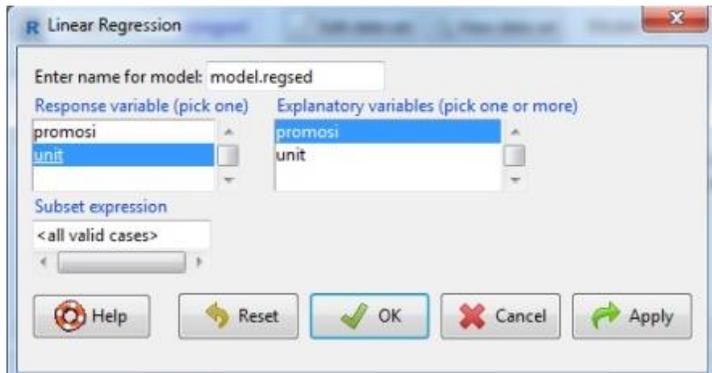
	unit	promosi
1	123	125.6
2	124	128.9
3	127	132.1
4	135	145.4
5	124	124.5
6	132	136.5
7	130	138.0
8	138	142.1
9	154	178.9
10	142	156.7
11	124	126.8
12	134	159.8
13	125	124.0
14	128	136.5
15	120	124.5
16	142	172.1
17	140	162.4
18	147	154.9
19	148	189.0
20	139	154.2
21	139	156.8
22	132	146.8
23	128	135.6
24	146	178.4

Kemudian akan dilakukan analisis regresi sederhana dengan langkah,

Klik menu statistics → fit models → linear regression...



Masukkan nama untuk model regresi misalkan model.regsed, dengan menjadikan jumlah produk yang dijual (unit) sebagai response variable dan variabel biaya promosi sebagai variabel ekplanatori.



Ouput yang diperoleh adalah sebagai berikut :

```
> model.regsed <- lm(unit~promosi,
data=latregsed)
> summary(model.regsed)
Call:
lm(formula = unit ~ promosi, data = latregsed)
Residuals:
    Min       1Q   Median       3Q      Max
-5.8693 -2.0773 -0.7734  1.5706  9.3155
```

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 68.61541    5.81822   11.79 5.55e-11
***
promosi      0.44589     0.03923   11.37 1.12e-10
***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
                '.' 0.1 ' ' 1

Residual standard error: 3.621 on 22 degrees of
freedom
Multiple R-squared:  0.8545, Adjusted  R-squared:
0.8479
F-statistic: 129.2 on 1 and 22 DF,  p-value:
1.12e-10

```

Dengan bentuk hipotesis yaitu :

$H_0$  = variabel unit jumlah produk yang terjual tidak berpengaruh terhadap variabel promosi

$H_1$  = variabel unit jumlah produk yang terjual berpengaruh terhadap variabel promosi

Dari output program R diatas diperoleh bahwa nilai statistika  $F = 129,2$  dengan  $p\text{-value } 1,12 \cdot 10^{-10}$ . Dengan demikian nilai  $p\text{-value}$  lebih kecil dari nilai tingkat kemaknaan yaitu  $0,05$ , sehingga dapat diambil keputusan bahwa  $H_0$  uji ditolak. Hal ini berarti variabel unit jumlah produk yang terjual berpengaruh terhadap variabel promosi.

Selanjutnya hasil uji parsial dapat dilihat pada Coefficients. Terlihat bahwa nilai signifikansi pengujian parameter-parameter regresi (yaitu pengujian

$\beta_0 = 0$  dan  $\beta_1 = 0$  sangat kecil yaitu  $0,000$ . Karena nilai signifikansi pengujian lebih kecil dari nilai tingkat kemaknaan yaitu  $0,05$  yang digunakan maka kedua hipotesis nol uji ini ditolak.

Oleh karena itu model persamaan yang menyatakan hubungan antara banyaknya suatu unit yang terjual dengan besarnya promosi yang dilakukan, yaitu  $Y = 68,61541 + 0,44589X$

Model tersebut dapat digunakan untuk mengestimasi atau menduga banyaknya unit yang terjual berdasarkan promosi yang dilakukan yaitu:

$$\hat{Y} = 68,61541 + 0,44589(200) = 155,815 \text{ atau } 156 \text{ unit}$$

Hal ini berarti bahwa rata-rata banyaknya unit tersebut dalam biaya promosi sebesar 200jt adalah 156 unit.

Hal lain perlu diperhatikan adalah nilai  $R^2$  (koefisien determinasi). Nilai  $R^2$  sebesar 0,8545 menyatakan bahwa 85,45% keragaman data (keragaman unit terjual) disebabkan oleh faktor besarnya promosi yang dilakukan, sedangkan sisanya merupakan keragaman alami yang diperoleh dari pembeli.

## 2. Analisis Regresi Berganda

Model persamaan regresi linier berganda dari suatu populasi adalah :

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon_i$$

$X_1, X_2, \dots, X_n$  adalah himpunan variabel bebas (*independent*).

$\alpha$  dan  $\beta_i$  adalah parameter yang nilainya tidak diketahui sehingga diduga dengan menggunakan statistik sampel.

$\varepsilon_i$  adalah komponen sisaan yang tidak diketahui nilainya (acak).

Model populasi regresi berganda ini disuga dengan menggunakan metode kuadrat terkecil (*Least Square Method*). Prinsip metode kuadrat terkecil adalah meminimumkan selisih kuadrat antara Y-pengamatan dan Y-dugaan. Sehingga model regresi linier berganda menjadi sebagai berikut :

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

dengan,

Y = variabel tak bebas

X<sub>i</sub> = variabel bebas

a = penduga bagi parameter  $\alpha$  *intercept* (titik potong)

b<sub>i</sub> = penduga bagi  $\beta_i$

Untuk menentukan statistika a dan b digunakan metode kuadrat terkecil dengan cara meminimumkan kuadrat sisaan

$$JKG = \sum_{i=1}^n e_1^2 = \sum_{i=1}^n (y_1 - a - bx_1)^2$$

sehingga diperoleh :

$$a = \bar{y} - b\bar{x}$$

dan

$$b = \frac{n \sum_{i=1}^n x_1 y_1 - \left( \sum_{i=1}^n x_1 \right) \left( \sum_{i=1}^n y_1 \right)}{n \sum_{i=1}^n x_1^2 - \left( \sum_{i=1}^n x_1 \right)^2}$$

### a) Jenis Pengujian Hipotesis

Pengujian hipotesis mengenai pendugaan statistik terhadap parameter terbagi menjadi dua macam, yaitu :

#### 1. Pengujian Hipotesis Satu Arah ( *One Ttail Ttest* )

##### a. Hipotesis Arah Kanan

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu > \mu_0$$

Disebut pengujian satu arah atas ( *upper tail test* ), menggunakan sebelah kanan dari kurva normal. Kesimpulannya :

H<sub>0</sub> ditolak apabila z<sub>hitung</sub> > z<sub>α</sub> dan H<sub>0</sub> diterima apabila z<sub>hitung</sub> ≤ z<sub>α</sub>.

##### b. Hipotesis Arah Kiri

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0$$

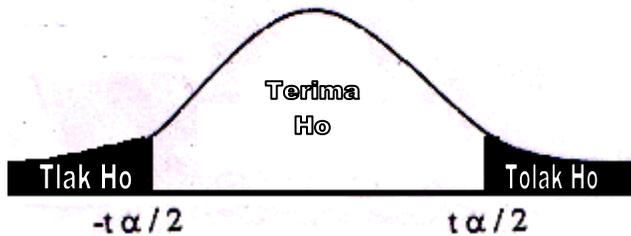
Disebut pengujian satu arah bawah (*lower tail test*), menggunakan sebelah kiri dari kurva normal. Kesimpulannya  $H_0$  ditolak apabila  $z_{\text{hitung}} < -z_{\alpha}$  dan  $H_0$  diterima apabila  $z_{\text{hitung}} \geq -z_{\alpha}$ .

## 2. Pengujian Hipotesis Dua Arah ( *Two Tail Test* )

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Disebut pengujian dua arah karena menggunakan daerah kritisnya terletak di sebelah kanan dan kiri kurva normal. Kesimpulannya  $H_0$  ditolak apabila  $z_{\text{hitung}} > z_{\alpha/2}$  atau  $z_0 < -z_{\alpha/2}$  dan  $H_0$  diterima apabila  $-z_{\alpha/2} \leq z_0 \leq z_{\alpha/2}$ . Gambar Daerah Penerimaan Dan Penolakan  $H_0$  adalah sebagai berikut :



### b) Pengujian Dalam Regresi Linier

Adapun pengujian terhadap persamaan regresi linear adalah sebagai berikut :

#### 1. Uji Kecukupan Data

Dari hasil output data dapat diperlihatkan hasil dari *p-value* lebih kecil dari nilai  $\alpha = 0,05$ , maka dari data tersebut dapat digunakan suatu analisis regresi sehingga tidak diperlukan adanya penambahan sampel. Apabila *p-value* lebih besar daripada nilai  $\alpha$ , maka dari data tersebut tidak dapat digunakan dalam suatu analisis regresi sehingga diperlukan adanya penambahan sampel. Uji kecukupan data diperlukan sekali sebelum melakukan regresi linier.

#### 2. Uji Koefisien Determinasi

Dari data output dapat diperlihatkan nilai dari  $R^2$ , hal ini membuktikan bahwa kekuatan dalam menjelaskan perubahan atau pergerakan variabel X terhadap variabel Y. Sedangkan  $R^2$

(adjusted) merupakan koefisien determinasi yang menjelaskan banyaknya variabel bebas dari data tersebut. Nilai  $R^2$  dinyatakan dengan rumus sebagai berikut :

$$R^2 = \frac{JK_{regresi}}{JK_{total}} = \frac{b'X'y - ny^2}{y'y - ny^2}$$

$R^2_{adjusted}$  dinyatakan dengan rumus sebagai berikut :

$$R^2_{adjusted} = 1 - (1 - R^2) \left( \frac{n-1}{n-k} \right)$$

### 3. Uji Koefisien Regresi (Parameter)

Uji ini juga tidak kalah pentingnya dengan uji-uji yang lain. Dari data di atas dapat digunakan pengujian dua cara yaitu pengujian serentak dengan menggunakan  $f$  dan pengujian parsial atau individu dengan  $t$ .

- Pengujian Secara Serentak (Dengan  $f$ )

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 : \beta_1 = \beta_2 = \dots = \beta_k \neq 0 \text{ atau } \beta_j \neq 0, \quad j = 1, 2, \dots, k$$

Dalam hal ini digunakan statistik uji ' $f$ ' yang dirumuskan sebagai berikut :

$$F_{hitung} = \frac{RK_{Reg}}{RK_{Res}} = \frac{(b'X'y - ny^2)/k}{(y'y - b'X'y)/(n-k-1)}$$

dengan derajat bebas regresi  $k=1$ , derajat bebas pembilang ( $v_1$ ) sama dengan  $n-1$  dan derajat bebas penyebut ( $v_2$ ) sama dengan  $n-k-1$ .

Apabila :

- b.  $f_{hitung} > f_{(\alpha, n-k-1)}$  maka  $H_0$  ditolak, artinya model persamaan regresi benar.
- b.  $f_{hitung} \leq f_{(\alpha, n-k-1)}$  maka gagal tolak  $H_0$ , artinya model persamaan regresi salah.

- Pengujian Secara Individu (Dengan t)

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, \quad j = 0, 1, 2, \dots, k$$

Rumus atau statistik uji yang digunakan adalah :

$$|t_{hit}| = \frac{b_j}{S(b_j)}$$

Apabila :

- $t_{hitung} > t_{(\alpha/2, n-k-1)}$  maka  $H_0$  ditolak, artinya ada pengaruh  $X_i$  yang sesuai dengan koefesien regresi terhadap model.
- $t_{hitung} \leq t_{(\alpha/2, n-k-1)}$  maka gagal tolak  $H_0$ , artinya tidak ada pengaruh  $X_i$  yang sesuai dengan koefesien regresi terhadap model.

#### 4. Uji Kenormalan Data

Dari *output* data, untuk mengetahui kenormalan suatu data agar bisa digunakan dalam regresi linear digunakan pengujian sebagai berikut :

$H_0$  : Data berdistribusi normal

$H_1$  : Data tidak berdistribusi normal

Dimana harus diketahui nilai *p-value* dan  $\alpha$  terlebih dahulu.

Dari *output* data diketahui bahwa *p-value* lebih kecil dari  $\alpha$  (*p-value* <  $\alpha$ ) maka hipotesis nol ( $H_0$  : data berdistribusi normal ) ditolak, apabila dari *output* data diketahui bahwa *p-value* lebih besar dari  $\alpha$  (*p-value* >  $\alpha$  ) maka hipotesis nol ( $H_0$  : data tidak berdistribusi normal) gagal ditolak.

#### c) *P-Value*

Nilai *p-value* merupakan suatu probabilitas dalam menemukan uji statistik lebih besar dari yang dihasilkan. Apabila *p-value* lebih kecil dari nilai  $z_0$  maka  $H_0$  ditolak, jika *p-value* lebih besar dari nilai  $z_0$  maka  $H_0$  diterima. Nilai *p-value* yang sangat kecil contohnya 0.0001 menunjukkan probabilitas  $H_0$  benar

sangat kecil. Namun, di sisi lain  $p$ -value 0.233 berarti  $H_0$  gagal ditolak dan sangat besar kemungkinan benar. Dengan  $p$ -value kita dapat menunjukkan kemungkinan  $H_0$  itu benar atau salah. Apabila nilai  $p$ -value semakin besar, maka semakin besar pula kemungkinan  $H_0$  diterima.

Contoh kasus :

Seorang peneliti melakukan pengamatan terhadap tingkat keuntungan, tingkat penerimaan bersih untuk tiap unit rupiah dan jumlah kantor pada sebuah lembaga keuangan selama 25 tahun pengamatan.

Penyelesaian :

Hipotesis yang diuji adalah :

$H_0$  adalah variabel independen tidak secara bersama-sama berpengaruh terhadap variabel respon

$H_1$  adalah variabel independen berpengaruh secara bersama-sama terhadap variabel respon

Untuk menguji hipotesis ini digunakan uji serempak (uji F).

$H_0 = \beta_0 = \beta_1 = \beta_2 = 0$  (tidak ada pengaruh secara persial antara tingkat keuntungan dan tingkat penerimaan bersih terhadap jumlah kantor pada sebuah lembaga keuangan).

$H_1 =$  minimal ada satu  $\beta \neq 0$  (terdapat pengaruh secara persial antara tingkat keuntungan dan tingkat penerimaan bersih terhadap jumlah kantor pada sebuah lembaga keuangan).

Untuk menguji hipotesis ini digunakan uji parsial (uji t).

Prosedur uji regresi linier berganda menggunakan R-Commander adalah:

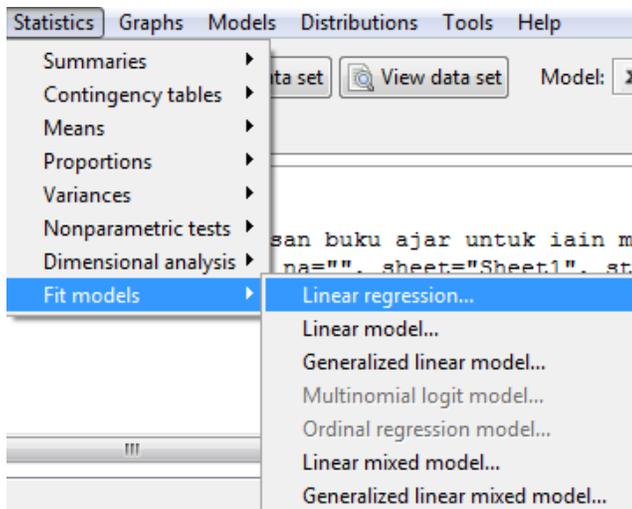
Melakukan input data dengan fasilitas impor data yang sudah diinput didalam Excel dengan klik menu Data  $\rightarrow$  impor data  $\rightarrow$  from excel file...

Masukkan nama untuk dataset misalkan latregganda, klik view data set untuk memastikan data yang kita import benar.

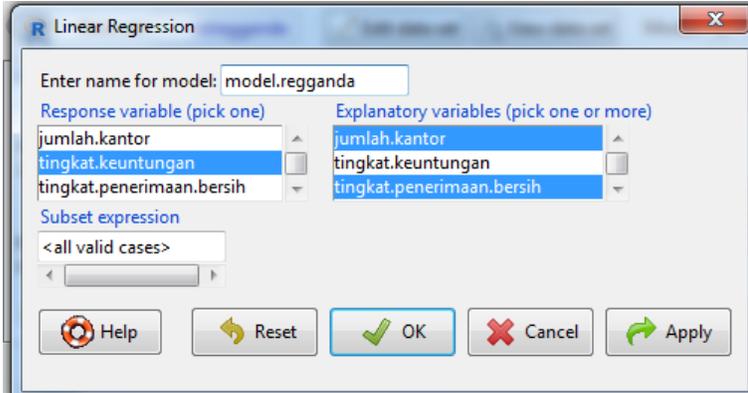
	tingkat.keuntungan	tingkat.penerimaan.bersih	jumlah.kantor
1	0.75	3.92	7298
2	0.71	3.61	6855
3	0.66	3.32	6636
4	0.61	3.07	6506
5	0.70	3.06	6450
6	0.72	3.11	6402
7	0.77	3.21	6368
8	0.74	3.26	6340
9	0.90	3.42	6349
10	0.82	3.42	6352
11	0.75	3.45	6361
12	0.77	3.58	6369
13	0.78	3.66	6546
14	0.84	3.78	6672
15	0.79	3.82	6890
16	0.70	3.97	7115
17	0.68	4.07	7327
18	0.72	4.25	7546
19	0.55	4.41	7931
20	0.63	4.49	8097
21	0.56	4.70	8468
22	0.41	4.58	8717
23	0.51	4.69	8991
24	0.47	4.71	9179
25	0.32	4.78	9318

Kemudian akan dilakukan analisis regresi sederhana dengan langkah,

Klik menu statistics → fit models → linear regression...



Masukkan nama untuk model regresi berganda misalkan model.regganda, dengan menjadikan tingkat keuntungan sebagai response variable dan jumlah kantor dan tingkat penerimaan bersih sebagai variabel ekplonatori.



Ouput yang diperoleh adalah sebagai berikut :

```
>                                model.regganda                                <-
  lm(tingkat.keuntungan~jumlah.kantor+tingkat.penerimaa
    n.bersih, data=latregganda)
> summary(model.regganda)
Call:
lm(formula = tingkat.keuntungan ~ jumlah.kantor +
    tingkat.penerimaan.bersih,
    data = latregganda)
Residuals:
    Min       1Q   Median       3Q      Max
-0.085090 -0.039102 -0.003341  0.030236  0.105692
Coefficients:
                                Estimate   Std. Error  t
value Pr(>|t|)
(Intercept)                   1.56449677   0.07939598
 19.705 1.82e-15 ***
jumlah.kantor                  -0.00024908   0.00003205  -
 7.772 9.51e-08 ***
tingkat.penerimaan.bersih      0.23719747   0.05555937
 4.269 0.000313 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1
```

Residual standard error: 0.0533 on 22 degrees of freedom  
Multiple R-squared: 0.8653, Adjusted R-squared:  
0.8531

F-statistic: 70.66 on 2 and 22 DF, p-value: 2.65e-10

Berdasarkan hasil output diatas dapat dijelaskan bahwa untuk hasil uji serempak (uji F) diperoleh bahwa nilai statistic F = 70,66 dengan p-value =  $2,65 \cdot 10^{-10}$ . Dengan demikian p-value lebih kecil dari tingkat kemaknaan yaitu 0,05, sehingga dapat diambil keputusan bahwa  $H_0$  uji tersebut ditolak. Hal ini berarti bahwa jumlah kantor dan tingkat penerimaan berpengaruh terhadap tingkat keuntungan.

Selanjutnya, hasil uji parsial (uji t) dapat dilihat pada tabel Coefficients. Terlihat bahwa nilai signifikansi pengujian parameter-parameter regresi yaitu pengujian  $\beta_0 = 0, \beta_1 = 0$  dan  $\beta_3 = 0$  sangat kecil yaitu 0,000. Karena nilai signifikansi pengujian lebih kecil dari nilai tingkat kemaknaan yaitu 0,05 yang digunakan, maka ketiga hipotesis nol uji ini ditolak. Hal ini berarti bahwa  $\beta_0, \beta_1$  dan  $\beta_2 \neq 0$  atau terdapat pengaruh secara parsial jumlah kantor dan tingkat penerimaan berpengaruh terhadap tingkat keuntungan. Oleh sebab itu, model persamaan yang menyatakan pengaruh antara jumlah kantor dan tingkat penerimaan berpengaruh terhadap tingkat keuntungan Adalah sebagai berikut :

$$\hat{Y} = 1,56449677 - 0,00024908X_1 + 0,23719747X_2$$

Model tersebut dapat digunakan untuk mengestimasi atau menduga dimana masing-masing parameter yang diestimasi memiliki interpretasi yang berbeda yaitu :

a = 1,56449677 ini berarti apabila jumlah kantor dan tingkat penerimaan sebesar nol, maka nilai y akan sebesar 1,565

b<sub>1</sub> = -0,000249, ini berarti apabila ada kenaikan tingkat penerimaan sebesar 1 unit sedangkan nilai variable ekplonatori lainnya tetap, maka nilai y diharapkan turun sebesar 0,000249

$b_2 = 0,237$ , ini berarti apabila ada kenaikan jumlah kantor sebesar 1 unit sedangkan nilai variable eksplonatori lainnya tetap, maka nilai  $y$  diharapkan naik 0,237.

Berdasarkan hasil uji serempak (uji F) dapat disimpulkan bahwa secara bersama-sama jumlah kantor dan tingkat penerimaan berpengaruh secara signifikan terhadap tingkat keuntungan. Hal ini dapat dilihat dari nilai p-value uji F sebesar  $2,65 \cdot 10^{16}$  dimana nilainya jelas lebih kecil dari tingkat kenyataan yang digunakan yaitu 0,05. Berdasarkan uji t (uji parsial) dapat disimpulkan bahwa tingkat penerimaan bersih untuk tiap unit rupiah berpengaruh secara signifikan terhadap tingkat keuntungan, hal ini dapat dilihat dari p-value uji t untuk tingkat penerimaan pada kolom Coeffisients sebesar 0.000313 lebih kecil dari taraf signifikansi yaitu 0,05. Demikian juga jumlah kantor berpengaruh secara signifikan terhadap tingkat keuntungan, hal ini dapat dilihat dari p-value Uji-t untuk jumlah kantor pada kolom Coeffisients sebesar  $9,5 \cdot 10^{-8}$ .

Perlu diperhatikan juga adalah nilai  $R^2$  (koefisien determinasi). Nilai  $R^2$  sebesar 0,8653 menyatakan bahwa 86,53 % keragaman data (keragaman tingkat keuntungan) disebabkan karena factor tingkat penerimaan bersih untuk tiap unit rupiah dan jumlah kantor, sedangkan yang lain sisannya meruoakan keragaman alami data yang diperoleh.

### **C. Rangkuman**

Analisis regresi adalah teknik statistika yang berguna untuk memeriksa dan memodelkan hubungan diantara variabel-variabel. Penerapannya dapat dijumpai secara luas di banyak bidang seperti teknik, ekonomi, manajemen, ilmu-ilmu biologi, ilmu-ilmu sosial, dan ilmu-ilmu pertanian. Pada saat ini, analisis regresi berguna dalam menelaah hubungan dua variabel atau lebih, dan terutama untuk menelusuri pola

hubungan yang modelnya belum diketahui dengan sempurna, sehingga dalam penerapannya lebih bersifat eksploratif.

Analisis regresi dikelompokkan dari mulai yang paling sederhana sampai yang paling rumit, tergantung tujuan yang berlandaskan pengetahuan atau teori sementara, bukan asal ditentukan saja. Adapun jenisnya ada empat antara lain Regresi linier sederhana bertujuan mempelajari hubungan linier antara dua variabel. Dua variabel ini dibedakan menjadi variabel bebas ( $X$ ) dan variabel tak bebas ( $Y$ ). Variabel bebas adalah variabel yang bisa dikontrol sedangkan variabel tak bebas adalah variabel yang mencerminkan respon dari variabel bebas.

Dan Regresi berganda seringkali digunakan untuk mengatasi permasalahan analisis regresi yang melibatkan hubungan dari dua atau lebih variabel bebas. Pada awalnya regresi berganda dikembangkan oleh ahli ekonometri untuk membantu meramalkan akibat dari aktivitas-aktivitas ekonomi pada berbagai segmen ekonomi. Misalnya laporan tentang peramalan masa depan perekonomian di jurnal-jurnal ekonomi (*Business Week*, *Wal Street Journal*, dll), yang didasarkan pada model-model ekonometrik dengan analisis berganda sebagai alatnya.

#### **D. Tugas**

Lakukan analisis regresi sederhana dan analisis regresi berganda dari hasil pencarian data yang sudah dilakukan, kemudian berikan analisis dan interpretasi dari hasil pengujian anda.

### **E. Rujukan**

- Aridinanti, L. dkk. (2003). *Pengantar Metoda Statistika*. Surabaya : Jurusan Statistika FMIPA ITS
- Aunuddin. (1989). *Analisis Data*. Bogor: Institut Pertanian Bogor
- Damanhuri, Enri. (2001). *Statistika Lingkungan*. Bandung: Departemen Teknik Lingkungan ITB
- Riduwan. (2004). *Metode dan Teknik Menyusun Tesis*. Bandung: Penerbit Alfabeta.
- Rumiati, A.T. (1998). *Konsep Peluang dan Statistika dalam Rekayasa*. Surabaya : Jurusan Statistika FMIPA ITS
- Siagian, D. dan Sugiarto. (2002). *Metode Statistika Untuk Bisnis Dan Ekonomi*. Jakarta : PT. Gramedia Pustaka Utama
- Sumargo, Chr H. (1984). *Pendahuluan Teori Kemungkinan dan Statistika*. Bandung : ITB
- Walpole, E. R. (1995). *Pengantar Statistik Edisi Ketiga*. Jakarta : PT. Gramedia Pustaka Utama
- \_\_\_\_\_.Desember (2005). *Statistik Ekonomi Keuangan Indonesia*: Bank Indonesia

### **F. Bacaan yang Dianjurkan**

- Aridinanti, L. dkk. (2003). *Pengantar Metoda Statistika*. Surabaya : Jurusan Statistika FMIPA ITS
- Aunuddin. (1989). *Analisis Data*. Bogor: Institut Pertanian Bogor
- Damanhuri, Enri. (2001). *Statistika Lingkungan*. Bandung : Departemen Teknik Lingkungan ITB
- Riduwan. (2004). *Metode dan Teknik Menyusun Tesis*. Bandung: Penerbit Alfabeta.
- Rumiati, A.T. (1998). *Konsep Peluang dan Statistika dalam Rekayasa*. Surabaya : Jurusan Statistika FMIPA ITS
- Siagian, D. dan Sugiarto. (2002). *Metode Statistika Untuk Bisnis Dan Ekonomi*. Jakarta : PT. Gramedia Pustaka Utama

- Sumargo, Chr H. (1984). *Pendahuluan Teori Kemungkinan dan Statistika*. Bandung : ITB
- Walpole, E. R. (1995). *Pengantar Statistik Edisi Ketiga*. Jakarta : PT. Gramedia Pustaka Utama
- \_\_\_\_\_.Desember (2005). *Statistik Ekonomi Keuangan Indonesia*: Bank Indonesia

# **BAB XIII**

## **METODE DATA ENVELOPMENT ANALYSIS (DEA)**

### **A. Pendahuluan**

Pada Bab ini akan dibahas teori Metode Data Envelopment Analysis (DEA), serta aplikasinya dalam persoalan Ekonomi dan Bisnis menggunakan bantuan program R.

#### **Tujuan Pembelajaran :**

Setelah mempelajari materi pada bab ini mahasiswa diharapkan menerapkan teori Metode Data Envelopment Analysis (DEA), serta aplikasinya dalam persoalan Ekonomi dan Bisnis menggunakan bantuan program R.

#### **Capaian Pembelajaran :**

1. Sikap dan tata nila  
CPMK 1; Mahasiswa diharapkan mampu menerapkan Statistik Bisnis (PAKEM). (P5)
2. Pengetahuan  
CPMK 2; Mahasiswa mampu melakukan analisis Statistik Bisnis menggunakan bantuan Software). (P 4)
3. Keterampilan  
CPMK 3; Mahasiswa dapat menginterpretasikan Analisis metode Statistika dalam bidang Ekonomi dan Bisnis dan menarik kesimpulan. (S12)

#### **Materi**

1. Teori Metode Data Envelopment Analysis (DEA)
2. Model Constant Return to Scale (CRS)
3. Model Variabel Return to Scale (VRS)
4. Menerapkan DEA dengan bantuan Program DEAP

## B. Teori Metode Data Envelopment Analysis (Dea)

Data Envelopment Analysis (DEA) merupakan sebuah metode optimasi program matematika yang mengukur efisiensi teknik suatu Decision Making Unit (DMU), dan membandingkan secara relatif terhadap DMU yang lain. Teknik analisis DEA didesain khusus untuk mengukur efisiensi relatif suatu DMU dalam kondisi banyak input maupun output. Efisiensi relatif suatu DMU adalah efisiensi suatu DMU dibanding dengan DMU lain dalam sampel yang menggunakan jenis input dan output yang sama. DEA memformulasikan DMU sebagai program linear fraksional untuk mencari solusi, apabila model tersebut ditransformasikan ke dalam program linear dengan nilai bobot dari input dan output. Efisiensi relatif DMU dalam DEA juga didefinisikan sebagai rasio dari total output tertimbang dibagi total input tertimbang (total weighted output/total weighted input). Setiap DMU diasumsikan bebas menentukan bobot untuk setiap

Variabel-variabel input maupun output yang ada, asalkan mampu memenuhi dua kondisi yang disyaratkan, yakni:

1. Bobot tidak boleh negatif
2. Bobot harus bersifat universal.

Hal ini berarti setiap DMU dalam sampel harus dapat menggunakan seperangkat bobot yang sama untuk mengevaluasi rasionya (total weighted output/total weighted input) dan rasio tersebut tidak lebih dari 1 (total weighted output/total weighted input  $\leq 1$ ).

DEA berasumsi bahwa setiap DMU akan memiliki bobot yang memaksimalkan rasio efisiensinya (maximize total weighted output/total weighted input). Asumsi maksimisasi rasio efisiensi ini menjadikan penelitian DEA ini menggunakan orientasi output dalam menghitung efisiensi

teknik. Orientasi lainnya adalah meminimalisasi input, namun kedua asumsi tersebut akan diperoleh hasil yang sama.

Suatu DMU dikatakan efisien secara relatif apabila nilai dualnya sama dengan 1 (nilai efisiensi 100 persen), sebaliknya apabila nilai dualnya kurang dari 1 maka DMU bersangkutan dianggap tidak efisien secara relatif atau mengalami inefisiensi.

Efisiensi adalah kemampuan untuk menyelesaikan suatu pekerjaan dengan benar atau dalam pandangan matematika didefinisikan sebagai perhitungan rasio output (keluaran) dan atau input (masuk) atau jumlah keluaran yang dihasilkan dari suatu input yang digunakan. Dalam Kamus Besar Bahasa Indonesia, efisiensi diterjemahkan dengan daya guna. Ini menunjukkan bahwa efisiensi selain menekankan pada hasilnya, juga ditekankan pada daya atau usaha/pengorbanan untuk mencapai hasil tersebut agar tidak terjadi pemborosan.

*Decision Making Unit* (DMU) merupakan istilah yang digunakan terhadap unit yang akan diukur efisiensinya. Dalam hal ini, penelitian dengan pendekatan DEA akan menganalisis efisiensi relatif suatu DMU dalam satu kelompok observasi terhadap DMU lain dengan kinerja terbaik dalam kelompok observasi tersebut. Ada beberapa hal yang dianggap penting untuk diperhatikan dalam pemilihan DMU dan variabel input-output antara lain:

1. *Positivity*

DEA menuntut semua variabel input atau output bernilai positif.

2. *Isotonicity*

Variabel input dan output harus memiliki hubungan isotonicity yang berarti untuk setiap kenaikan pada variabel input apapun harus menghasilkan kenaikan

setidaknya satu variabel output dan tidak ada variabel output yang mengalami penurunan.

3. Jumlah DMU

Dibutuhkan setidaknya jumlah DMU sebesar 3 kali dari jumlah variabel input dan output.

4. Window analysis

Perlu dilakukan window analysis jika terjadi pemecahan data DMU (tahunan menjadi triwulan misalnya) yang biasanya dilakukan untuk memenuhi syarat jumlah DMU. Analisis ini dilakukan untuk menjamin stabilitas nilai efisiensi dari DMU yang bersifat time dependent.

5. Penentuan bobot

Walaupun DEA menentukan bobot yang seringan mungkin untuk setiap unit relatif terhadap unit yang lain dalam satu set data, terkadang dalam praktek manajemen dapat menentukan bobot sebelumnya.

6. Homogeneity

DEA menuntut seluruh DMU yang di evaluasi memiliki variabel input dan output yang sama jenisnya. Berdasarkan seluruh ketentuan tersebut, DMU yang dipilih dalam penelitian ini adalah perusahaan asuransi penjaminan.

**1. Model Constant Return to Scale (CRS)**

Model constant return to scale dikembangkan oleh Charnes, Cooper dan Rhodes (Model CCR) pada tahun 1978. Model ini mengasumsikan bahwa rasio antara penambahan input dan output adalah sama (constant return to scale). Artinya, jika ada tambahan input sebesar  $x$  kali, maka output akan meningkat sebesar  $x$  kali juga. Asumsi lain yang digunakan dalam model ini adalah bahwa setiap perusahaan atau Decision Making Unit (DMU) beroperasi pada skala yang optimal.

Rumus dari constant return to scale dapat dituliskan sebagai berikut:

Max  $\theta$  (Efisiensi DMU Model CRS)

$$\sum_j^n = 1 \quad x_{ij} \geq \theta x_{i0} \quad i = 1, 2, \dots, m$$

$$\sum_j^n = 1 \quad y_{rj} \geq \theta y_{r0} \quad r = 1, 2, \dots, s$$

$$\sum_j^n = 1 \quad \lambda_j \geq 0 \quad i = 1, 2, \dots, n$$

Di mana:

$\theta$  = efisiensi teknis (CRS)

n = jumlah DMU

m = jumlah input

s = jumlah output

$x_{ij}$  = jumlah input tipe ke-i dari DMU ke-j

$y_{rj}$  = jumlah output tipe ke-r dari DMU ke-j

$\lambda_j$  = bobot DMU j untuk DMU yang dihitung

Nilai efisiensi selalu kurang atau sama dengan 1. DMU yang nilai efisiensinya kurang dari 1 berarti inefisiensi sedangkan DMU yang nilai efisiensinya sama dengan 1 berarti DMU tersebut efisien.

## 2. Model variabel Return to Scale (VRS)

Model ini dikembangkan oleh Banker, Charnes, dan Cooper (model BCC) pada tahun 1984 dan merupakan pengembangan dari model CCR. Model ini beranggapan bahwa perusahaan tidak atau belum beroperasi pada skala yang optimal. Asumsi dari model ini adalah bahwa rasio antara penambahan input dan output tidak sama (variable return to scale). Artinya, penambahan input sebesar x kali tidak akan menyebabkan output meningkat sebesar x kali, bisa lebih kecil atau lebih besar dari x kali. Peningkatan proporsi bisa bersifat increasing return to scale (IRS) atau bisa juga bersifat decreasing return to scale (DRS). Hasil model ini menambahkan kondisi

convexity bagi nilai-nilai bobot, dengan memasukkan dalam model batasan berikut:

$$\sum_{j=1}^n x_j = 1$$

Selanjutnya model BCC dapat ditulis dengan persamaan berikut:

Max (Efisiensi DMU Model VRS)

$$\sum_j^n = 1 x_{ij} \quad 'ij \geq xi0 \quad i = 1,2,\dots,m$$

$$\sum_j^n = 1 y_{rj} \quad 'j \geq yi0 \quad r = 1,2,\dots,j$$

$$\sum_j^n = 1 \quad 'j \geq 1 \quad (VRS)$$

$$\sum_j^n = 1 \quad 'j \geq 0 \quad j = 1,2,\dots,n$$

Di mana:

$\theta$  = efisiensi teknis (VRS)

n = jumlah DMU

m = jumlah input

s = jumlah output

$x_{ij}$  = jumlah input tipe ke-i dari DMU ke-j

$y_{rj}$  = jumlah output tipe ke-r dari DMU ke-j

'j = bobot DMU j untuk DMU yang dihitung

Nilai dari efisiensi tersebut selalu kurang atau sama dengan 1. DMU yang nilai efisiensinya kurang dari 1 berarti inefisiensi sedangkan DMU yang nilainya sama dengan 1 berarti DMU tersebut efisien.

### 3. Implementasi metode Data Envelopment Analysis Menggunakan Program R

Berbeda dengan implementasi metode yang lain pada bab-bab sebelumnya untuk Metode Dea tidak bias digunakan menggunakan bantuan R-Commander karena menu belum tersedia di dalamnya. Akan tetapi masih bias di selesaikan menggunakan Program R khususnya

RStudio yaitu pengembangan dari Program R yang bersifat *command line* artinya harus mengetikkan perintah terlebih dahulu untuk menyelesaikan perhitungan dalam Metode DEA yaitu dengan menjalankan fungsi DEA yang digunakan dalam Metode Data Envelopment Analysis. Fasilitas yang berupa fungsi program tersebut merupakan implementasi dari penerapan dalam fungsi DEA. Sehingga para statistisi dan para pengguna pada umumnya yang ingin menganalisis fungsi DEA, dapat memanggil fungsi DEA tersebut untuk kemudian menjalankannya.

Berikut ini struktur fungsi DEA yang terdapat dalam pustaka `dear` pada Software R dan dapat digunakan dalam Metode DEA, adalah sebagai berikut :

DESKRIPSI :

Orientasi penyelesaian dari model yang mendasar dari metode DEA berbentuk input dan output di bawah konstan (model CCR DEA), variabel (model DEA BCC), skala pengembalian tidak meningkat, tidak menurun, atau digeneralisasi. Secara default, model diselesaikan dalam proses dua tahap (kelonggaran DEA dimaksimalkan).

Fungsi `model_basic` memungkinkan untuk menangani input/output yang tidak bijaksana, tidak dapat dikontrol, dan tidak diinginkan.

Akhirnya dapat digunakan fungsi `model_basic` untuk menyelesaikan metode DEA secara terarah dengan memilih `orientation = "dir"`.

PENGUNAAN :

```
model_basic(datadea,  
            dmu_eval = NULL,  
            dmu_ref = NULL,  
            orientation = c("io", "oo", "dir"),  
            dir_input = NULL,  
            dir_output = NULL,
```

```

rts = c("crs", "vrs", "nirs", "ndrs",
"grs"),
L = 1,
U = 1,
maxslack = TRUE,
weight_slack_i = 1,
weight_slack_o = 1,
vtrans_i = NULL,
vtrans_o = NULL,
compute_target = TRUE,
compute_multiplier = FALSE,
returnlp = FALSE,
...)

```

Keterangan lebih lanjut mengenai perintah-perintah di atas

<code>datadea</code>	Input data, memasukkan n DMU, m input dan s output.
<code>dmu_eval</code>	Vektor numerik yang berisi DMU yang harus dievaluasi. jika NULL(gagal), untuk semua DMU dipertimbangkan.
<code>dmu_ref</code>	Vektor numerik yang berisi DMU yang merupakan set referensi evaluasi. jika NULL (gagal), untuk semua DMU dipertimbangkan.
<code>orientation</code>	Sebuah rangkaian, kesamaan "io" (input oriented), "oo" (output oriented), or "dir" (directional).
<code>dir_input</code>	Sebuah nilai, panjang vektor m, atau matrik m x ne (dimana ne adalah panjang dari <code>dmu_eval</code> )

	dengan input terarah. jika <code>dir_input ==</code> matrik input (dari dalam DMU <code>dmu_eval</code> ) dan <code>dir_output == 0</code> , itu setara dengan berorientasi input ( $\beta = 1 - \text{efficiency}$ ). jika <code>dir_input</code> dihilangkan, matrik input (dari DMU dalam <code>dmu_eval</code> ) ditugaskan.
<code>dir_output</code>	Sebuah nilai, panjang vektor <code>s</code> , atau matrik <code>s x ne</code> (dimana <code>ne</code> adalah panjang dari <code>dmu_eval</code> ) dengan output terarah. jika <code>dir_input == 0</code> dan <code>dir_output ==</code> matrik output (dari dalam <code>dmu_eval</code> itu setara dengan berorientasi output ( $\beta = \text{efficiency} - 1$ ). jika <code>dir_output</code> dihilangkan, matrik output (dari DMU dalam <code>dmu_eval</code> ) ditugaskan.
<code>rts</code>	Sebuah rangkaian, menentukan jenis tipe pengembalian dari skala, sama untuk "crs" (constant), "vrs" (variable), "nirs" (non-increasing), "ndrs" (non-decreasing) atau "grs" (generalized).
<code>L</code>	Batas bawah dari generalized returns to scale (grs).

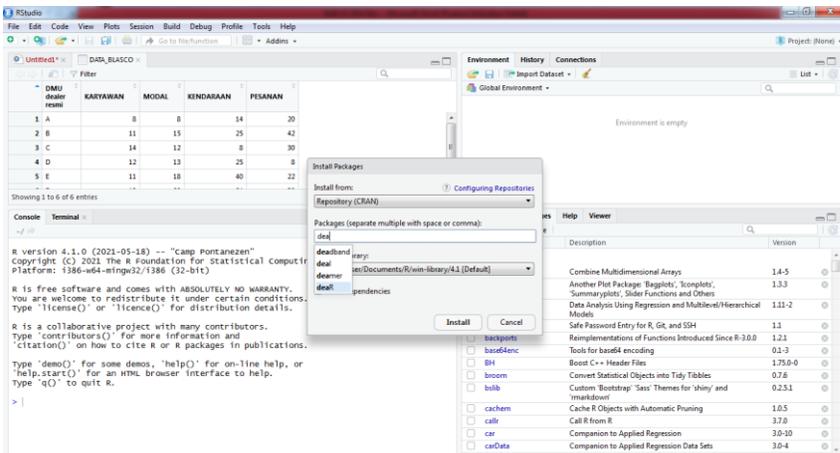
U	Batas atas generalized returns to scale (grs).
maxslack	Logika. jika TRUE, menghitung secara lambat untuk nilai maksimal.
weight_slack_i	Sebuah nilai, panjang vektor $m$ , atau matrik $m \times n_e$ (dimana $n_e$ adalah panjang dari <code>dmu_eval</code> ) dengan pembobot input untuk memaksimalkan solusi.
weight_slack_o	Sebuah nilai, panjang vektor $s$ , atau matrik $s \times n_e$ (dimana $n_e$ adalah panjang dari <code>dmu_eval</code> ) dengan pembobot output untuk memaksimalkan solusi.
vtrans_i	Vektor numerik terjemahan untuk input yang tidak diinginkan dengan orientasi non-arah. jika <code>vtrans_i[i]</code> adalah NA, maka berlaku " <code>max + 1</code> " menterjemahkan indeks ke $i$ input yang tidak diinginkan. jika <code>vtrans_i</code> adalah sebuah konstanta, maka diterapkan untuk menterjemahkan semua input yang tidak diinginkan. jika <code>vtrans_i</code> adalah NULL, maka diberlakukan " <code>max + 1</code> "

	menterjemahkan semua input yang tidak diinginkan.
<code>vtrans_o</code>	Vektor numerik terjemahan untuk output yang tidak diinginkan dengan orientasi non-arah, sejalan <code>vtrans_i</code> , tetapi diaplikasikan kepada output.
<code>compute_target</code>	Logika. jika TRUE, menghitung secara lambat untuk nilai maksimal.
<code>compute_multiplier</code>	Logika. Jika TRUE, menghitung berbagai (solusi ganda) maka orientation adalah "io" atau "oo".
<code>returnlp</code>	Logika. jika TRUE, maka mengembalikan masalah linier (fungsi tujuan untuk kendala) dari tahap 1
<code>...</code>	Diabaikan, karena masalah kompatibilitas.

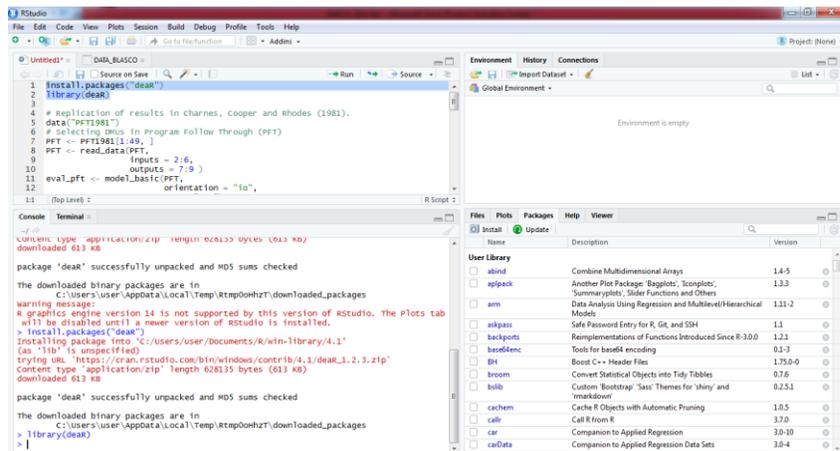
Contoh kasus : suatu peneliti ingin melakukan penelitian untuk menentukan efisiensi dari 6 DMU yaitu dealer resmi yang dimiliki suatu pabrik mobil, dengan dua input yaitu jumlah karyawan dan modal yang dimiliki dealer resmi tersebut dan dua output yaitu jumlah kendaraan dan order atau pesanan dari dealer tersebut.

Adapun langkah penyelesaian kasus tersebut menggunakan bantuan program R adalah

Langkah pertama instal terlebih dahulu paket deaR yang digunakan dalam menyelesaikan Metode DEA yang digunakan dalam penelitian tersebut dengan perintah `install.packages("deaR")`

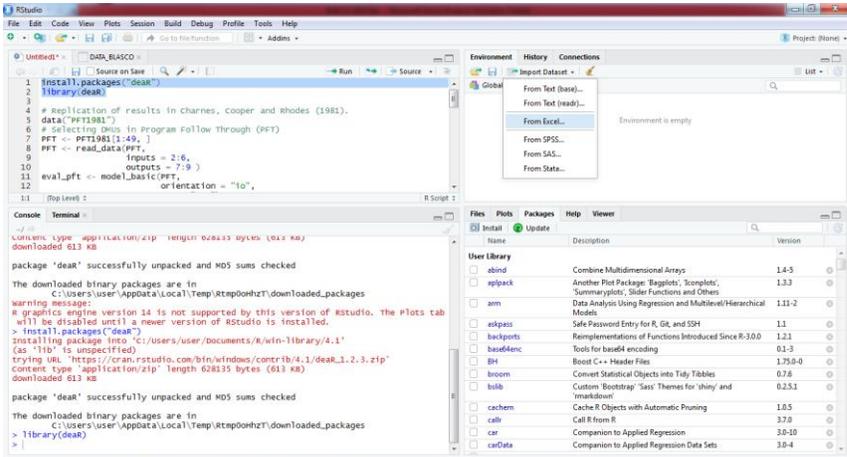


mengaktifkan paket deaR dengan perintah `library(deaR)`

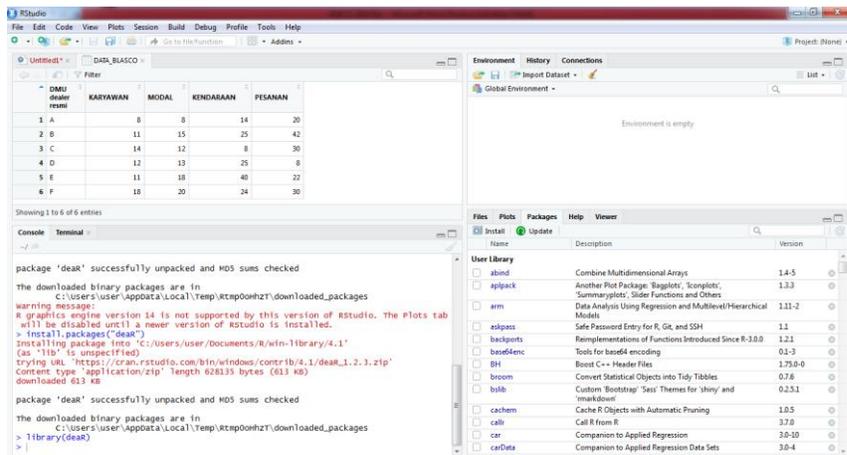


Kemudian klik Run

Input data data penelitian tersebut kedalam Excell kemudian lakukan import data dengan langkah seperti berikut



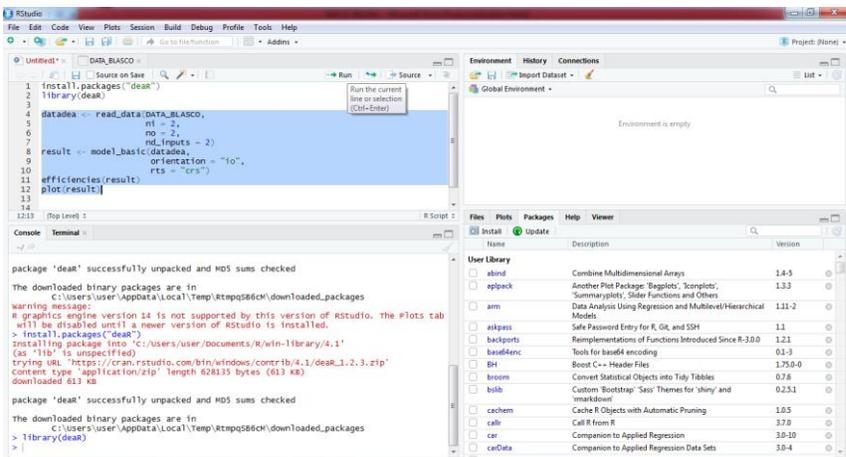
Berikut adalah data hasil inputnya :



menulis program Metode DEA yang digunakan untuk menyelesaikan kasus tersebut dengan perintah

```
datadea <- read_data(DATA_BLASCO,  
                    ni = 2,  
                    no = 2,  
                    nd_inputs = 2)  
result <- model_basic(datadea,  
                      orientation = "io",  
                      rts = "crs")  
efficiencies(result)  
plot(result)
```

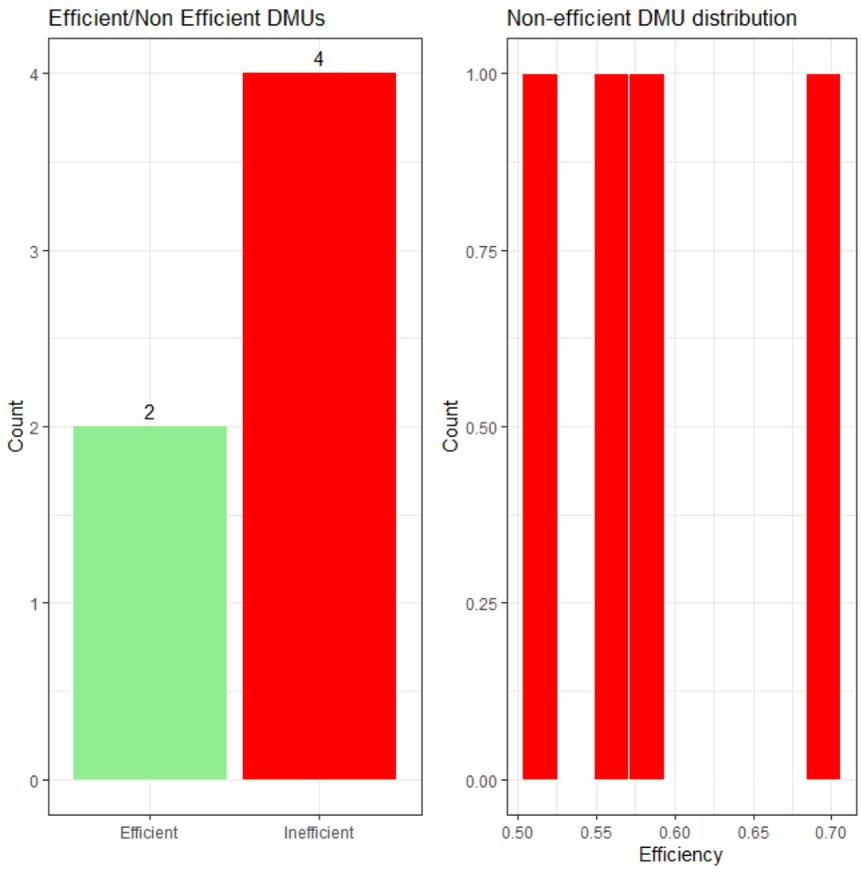
kemudian klik Run



Hasil dari output program tersebut adalah

```
> efficiencies(result)  
      A      B      C      D      E      F  
0.70575 1.00000 0.56122 0.57292 1.00000 0.50295  
> plot(result)
```

Press [enter] to continue



Dari hasil output diatas dapat dilihat bahwa berdasarkan hasil plot gambar terdapat DMU yang efisien sebanyak dua dealer resmi dan yang tidak efisien sebanyak enam buah dealer resmi, sedangkan nilai efisiensi diperoleh bahwa dari ke enam dealer resmi adalah :

Dealer Resmi (DMU)	Nilai Efisiensi
A	0,70575
B	1,00000
C	0,56122
D	0,57292
E	1,00000
F	0,50295

### C. Rangkuman

Data Envelopment Analysis (DEA) merupakan sebuah metode optimasi program matematika yang mengukur efisiensi teknik suatu Decision Making Unit (DMU), dan membandingkan secara relatif terhadap DMU yang lain. Teknik analisis DEA didesain khusus untuk mengukur efisiensi relatif suatu DMU dalam kondisi banyak input maupun output. Efisiensi relatif suatu DMU adalah efisiensi suatu DMU dibanding dengan DMU lain dalam sampel yang menggunakan jenis input dan output yang sama. DEA memformulasikan DMU sebagai program linear fraksional untuk mencari solusi, apabila model tersebut ditransformasikan ke dalam program linear dengan nilai bobot dari input dan output. Efisiensi relatif DMU dalam DEA juga didefinisikan sebagai rasio dari total output tertimbang dibagi total input tertimbang (total weighted output/total weighted input). Setiap DMU diasumsikan bebas menentukan bobot untuk setiap *Decision Making Unit* (DMU) merupakan istilah yang digunakan terhadap unit yang akan diukur efisiensinya.

Terdiri dari dua model yaitu Model constant return to scale dikembangkan oleh Charnes, Cooper dan Rhodes (Model CCR) pada tahun 1978. Model ini mengasumsikan bahwa rasio antara penambahan input dan output adalah sama (constant return to scale). Artinya, jika ada tambahan input sebesar  $x$  kali, maka output akan meningkat sebesar  $x$  kali juga dan Model ini dikembangkan oleh Banker, Charnes, dan Cooper (model BCC) pada tahun 1984 dan merupakan pengembangan dari model CCR. Model ini beranggapan bahwa perusahaan tidak atau belum beroperasi pada skala yang optimal. Asumsi dari model ini adalah bahwa rasio antara penambahan input dan output tidak sama (variable return to scale). Artinya, penambahan input sebesar  $x$  kali

tidak akan menyebabkan output meningkat sebesar  $x$  kali, bisa lebih kecil atau lebih besar dari  $x$  kali.

#### **D. Tugas**

Lakukan aplikasi penggunaan metode Data Envelopment Analysis (DEA) dari hasil pencarian data yang sudah dilakukan, kemudian berikan analisis dan interpretasi dari hasil pengujian anda.

#### **E. Rujukan**

Charnes, A.; Cooper, W.W.; Rhodes, E. (1978). "*Measuring the efficiency of decision making units*", European Journal of Operational Research 2, 429–444.

Charnes, A.; Cooper, W.W.; Rhodes, E. (1979). "*Short communication: Measuring the efficiency of decision making units*", European Journal of Operational Research 3, 339.

Charnes, A.; Cooper, W.W.; Rhodes, E. (1981). "*Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through*", Management Science, 27(6), 668-697.

Banker, R.; Charnes, A.; Cooper, W.W. (1984). "*Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis*", Management Science; 30; 1078-1092.

#### **F. Bacaan yang Dianjurkan**

Charnes, A.; Cooper, W.W.; Rhodes, E. (1978). "*Measuring the efficiency of decision making units*", European Journal of Operational Research 2, 429–444.

Charnes, A.; Cooper, W.W.; Rhodes, E. (1979). "*Short communication: Measuring the efficiency of decision*

making units”, *European Journal of Operational Research* 3, 339.

Charnes, A.; Cooper, W.W.; Rhodes, E. (1981). "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through", *Management Science*, 27(6), 668-697.

Banker, R.; Charnes, A.; Cooper, W.W. (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science*; 30; 1078-1092.

## DAFTAR PUSTAKA

- Allen, M.J and Yen, W.M., (1997), *Introduction to Measurement Theory*, Menterey: Books/ Cole Publishing Company.
- Alvin C, Rencher. (1998). *Multivariate Statistical Inference and Application*. Canada: John Willey & Sons.Inc.
- Aridinanti, L dkk. (2003). *Pengantar Metoda Statistika*. Surabaya: Jurusan Statistika FMIPA ITS.
- Aunuddin. (1989). *Analisis Data*. Bogor: Institut Pertanian Bogor
- Azwar,S., (2004), *Reliabilitas dan Validitas*, Pustaka Pelajar, Yogyakarta.
- Banker, R.; Charnes, A.; Cooper, W.W. (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science*; 30; 1078-1092.
- Charnes, A.; Cooper, W.W.; Rhodes, E. (1978). "Measuring the efficiency of decision making units", *European Journal of Operational Research* 2, 429-444.
- Charnes, A.; Cooper, W.W.; Rhodes, E. (1979). "Short communication: Measuring the efficiency of decision making units", *European Journal of Operational Research* 3, 339.
- Charnes, A.; Cooper, W.W.; Rhodes, E. (1981). "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through", *Management Science*, 27(6), 668-697.
- Cononer, W.J. (1980). *Practical Nonparametric Statistics Second Edition*. Canada: John Willey & Sons, Inc.
- Cononer, W. J. (1999). *Practical Nonparametric Statistic. Third Edition*. Texas, America : John Wiley & Sons, INC
- Cronbach, L.J, (1951), *Coefficient Alpha and The Internal Structure Of Test*, *Psychometrika*, 16, 297-334.

- Damanhuri, Enri. (2001). *Statistika Lingkungan*. Bandung : Departemen Teknik Lingkungan ITB
- Daniel, Wayne, W. (1989). *Statistik Nonparametrik Terapan*. Jakarta : PT Gramedia
- Dewi DAAN., (2018). Modul III: Uji Validitas dan Reliabilitas. *Statistika Terapan*. Universitas Diponegoro. 1-14.
- Djarwanto. (2003). *Statistik Nonparametrik*. Bpfe-Yogyakarta
- Furqon, (2001) *Statistika Terapan untuk Penelitian*, Bandung:Alfabeta
- Hadi, S (1995), *Statistik 1, 2, 3*, Yogyakarta: Andi Offset
- Hair,dkk.( 2006). *Multivariate Data Analysis sixth edition*. New Jersey: Pearson Plentice hall
- Iriawan, Nur dan Septin Puji A. (2006). *Mengolah Data Statistik dengan Mudah Menggunakan Minitab 14*. Yogyakarta : Andi
- Kuswardi dan Mutiara, Erna. (2004). *Statistik Berbasis Komputer untuk Orang-orang Nonstatistik*. Jakarta : PT Alex Media Komputindo
- Maindonald, J.H. (2004). *Using R for Data Analysis and Graphics: I ntroduction, Code and Commentary*. Centre for Bioinformat ion Science, Australian National University.
- Mason, R.D & Douglas A. Lind. (1996). *Teknik Statistik Untuk Bisnis dan Ekonomi*. Penerbit Erlangga, Jakarta.
- Nuryadi, dkk. (2017). [Dasar-dasar Statistik Penelitian](#). Yogyakarta: Gramasurya.
- Owen, W.J. (2007). *The R Guide*. Department of Mathematics and Computer Science, University of Richmond.
- Paradis, E. (2005). *R for Beginners*. Institut des Sciences de l' Evolution, Universite Montpellier II, France.
- Praptono. (1986). *Metode Statistika Nonparametrik*. Depdikbud UT.
- Purwanto. (2003). *Ekonomi Untuk Keuangan dan Bisnis*. Jakarta: Salemba Empat

- Riduwan. (2004). *Metode dan Teknik Menyusun Tesis*. Bandung : Penerbit Alfabeta.
- Rumiati, A.T. (1998). *Konsep Peluang dan Statistika dalam Rekayasa*. Surabaya : Jurusan Statistika FMIPA ITS
- Siagian, D. dan Sugiarto. (2002). *Metode Statistika Untuk Bisnis Dan Ekonomi*. Jakarta : PT. Gramedia Pustaka Utama
- Siegel, Sidney. (1986). *Statistik Nonparametrik untuk Ilmu-ilmu Sosial*. Jakarta : Gramedia
- Soepeno, Bambang. (1997). *Statistik Terapan dalam Penelitian Ilmu-ilmu Sosial dan Pendidikan*. Indonesia : PT Rineka Cipta
- Sudjana, (1992), *Metoda Statistika (Edisi ke 5)*, Bandung: Tarsito
- Sudjana. (1996). *Metoda Statistika (Edisi ke 6)*. Bandung : Tarsito
- Sugiyono. (2001). *Statistik Nonparametrik untuk Penelitian* . Cv Alfabeta Bandung.
- \_\_\_\_\_. (2003). *Statistika untuk Penelitian*. Bandung : CV Alfabeta
- \_\_\_\_\_. (2003). *Statistik Nonparametrik untuk Penelitian*. Bandung : CV Alfabeta
- Suhartono, (2008). *Analisis Data Statistik Dengan R*, Lab. Statistik Komputasi, ITS, Surabaya.
- Sumargo, Chr H. (1984). *Pendahuluan Teori Kemungkinan dan Statistika*. Bandung : ITB
- Supramono dan Sugiarto. (1993), *Statistika*, Yogyakarta: Andi Offset.
- Tirta, I. M. (2004). *Panduan R Pemrograman Untuk Analisis Data dan Grafik*. Jember:Laboratorium Statistika Jurusan Matematika FMIPA Universitas Jember.
- Usman, H. dan R. Purnomo Setiady Akbar. (2000). *Pengantar Statistika*. Jakarta : Bumi Aksara.
- Venables, W.N. and Smith, D.M. (2007). *An Introduction to R*. The R Development Core Team.
- Walpole, E R. (1995). *Pengantar Statistik Edisi Ketiga*. Jakarta : PT. Gramedia Pustaka Utama

- \_\_\_\_\_.Desember 2005. *Statistik Ekonomi Keuangan Indonesia*:  
Bank Indonesia
- Zanbar, Achmad S. (2005). Ilmu Statistika Pendekatan Teoritis  
dan Aplikatif disertai Contoh Penggunaan SPSS. Bandung :  
Rekayasa Sains.
- Zanten, V. W. (1980). Statistik untuk Ilmu- ilmu Sosial. Jakarta :  
Gramedia

## GLOSARI

### A

**Aksioma** adalah pernyataan yang dapat diterima sebagai kebenaran tanpa pembuktian.

**Analisis** adalah penyelidikan terhadap suatu peristiwa (karangan, perbuatan, dsb) untuk mengetahui keadaan yang sebenarnya (sebab-musabab, duduk perkaranya, dsb). Atau penguraian suatu pokok atas berbagai bagiannya dan penalaahan bagian itu sendiri serta hubungan antara bagian untuk memperoleh pengertian yang tepat dan pemahaman arti keseluruhan.

### B

**Berfikir Deduktif** adalah metode berpikir yang menerapkan hal-hal yang umum terlebih dahulu untuk seterusnya dihubungkan dalam bagian-bagiannya yang khusus.

**Berfikir Induktif** adalah metode berpikir yang menerapkan hal-hal khusus terlebih dahulu untuk seterusnya dihubungkan dalam bagian-bagian yang umum.

### D

**Data** adalah keterangan yang benar dan nyata atau keterangan atau bahan nyata yang dapat dijadikan dasar kajian (analisis atau kesimpulan)

**Datum** adalah yang sesuai dengan data.

**Data diskrit** adalah data yang didapatkan dengan cara menghitung atau membilang..

**Data kontinum** adalah data yang didapatkan dengan cara mengukur.

**Data kualitatif** adalah Data yang berbentuk kategori.

**Data kuantitatif** adalah Data yang berbentuk bilangan..

**Diagram** adalah gambaran (buram, seketsa) untuk memperlihatkan atau menerangkan sesuatu.

**Diagram Batang** adalah Digunakan untuk menyajikan data yang variabelnya berbentuk kategori atau atribut.

**Diagram gambar** adalah Digunakan untuk mendapatkan gambaran kasar sesuatu hal dengan menggunakan simbol. Setiap satuan jumlah diwakili oleh sebuah simbol sesuai dengan macam datanya.

**Diagram garis** adalah Menggambarkan data yang menerus atau berkesinambungan.

**Diagram lingkaran** adalah Menggambarkan proporsi masing – masing kategori data yang digambarkan dala satu lingkaran.

**Distribusi** adalah penyaluran (pembagian) kepada beberapa orang atau ke beberapa tempat.

**Distribusi Frekuensi** adalah suatu susunan data (organisasi data) statistik yang menunjukkan beberapa banyak hal dalam kategori-kategori atau interval yang berbeda dari data yang telah dikelompokkan.

## **E**

**Estimasi** adalah penilaian, pendapat, perkiraan, perhitungan, pangkal kulasian.

**Exhaustive (Lengkap terbatas)** yakni setiap individu, benda, atau pengukurannya harus muncul di satu kategori.

## **G**

**Grafik** adalah lukisan pasang surut suatu keadaan dengan garis atau gambar (tentang turun naiknya hasil, statistik dsb)

**Generalisasi** adalah perihal membentuk gagasan atau simpulan umum dari suatu kejadian, hal dsb.

## **H**

**Hipotesis** adalah sesuatu yang di anggap benar untuk alasan atau pengutaraan pendapat, meskipun kebenarannya masih harus dibuktikan; anggapan dasar.

**Histogram** adalah Merupakan diagram batang yang sisi-sisi berdekatnya berimpit..

**Hukum** adalah peraturan yang dibuat oleh penguasa (pemerintah) atau adat yang berlaku bagi semua orang disuatu masyarakat (Negara) atau undang-undang, peraturan, dan sebagainya untuk mengatur pergaulan hidup masyarakat.

## I

**Indikasi** adalah gejala keterkaitan masalah.

**Indikator** adalah alat pemantau (sesuatu) yang dapat memberikan petunjuk atau keterangan. Atau gejala yang menunjukkan keterkaitan.

## K

**Koefisien Variasi** adalah salah satu ukuran disperse relative yang membandingkan variasi antara nilai-nilai besar dengan nilai-nilai kecil dispersi relative.

**Konsep** adalah ide atau pengertian yang diabstrakkan dari peristiwa kongkret, atau gambaran mental dari obyek, proses atau apapun yang ada diluar bahasa yang digunakan oleh akal budi untuk memahami hal-hal lain.

**Konstruk** adalah sifat yang akan dipelajari.

## L

**Logika** adalah pengetahuan tentang kaidah berpikir atau jalan pikiran yang masuk akal.

## M

**Mean** adalah teknik penjelasan kelompok yang didasarkan atas nilai rata-rata dari kelompok tersebut.

**Median** adalah salah satu teknik penjelasan kelompok yang didasarkan atas nilai tengah dari kelompok data yang telah disusun urutannya dari yang terkecil sampai yang terbesar atau sebaliknya.

**Modus** adalah teknik penjelasan kelompok yang didasarkan atas nilai yang sedang populer atau nilai yang sering muncul dalam kelompok tersebut.

**Mutually Exclusive (Saling Lepas)** yakni suatu individu, benda atau pengukurannya hanya tercakup dalam satu kategori saja.

**Meramal (forecasting)** adalah menduga.

**Menafsirkan** adalah mengartikan.

## O

**Ogive** adalah suatu grafik yang mencantumkan frekuensinya secara meningkat dan menggunakan batas nyata.

## P

**Parameter** adalah Besaran, nilai, atau harga dari populasi

**Penalaran** adalah proses pemikiran secara logis untuk menarik kesimpulan dari suatu kenyataan sebelumnya.

**Populasi** adalah sekumpulan orang atau objek yang sedang diteliti

**Poligon frekuensi** adalah Jika titik – titik tengah sisi atas yang berdekatan dihubungkan dengan garis – garis patah, maka didapatkan suatu polygon frekuensi.

**Postulat** adalah asumsi yang menjadi pangkal dalil yang dianggap benar tanpa perlu membuktikannya; anggapan dasar; patokan duga.

**Proposisi** adalah rancangan usulan atau ungkapan yang dapat dipercaya, disangsikan, disangkal atau dibuktikan benar tidaknya.

**Poligon** adalah grafik garis yang menghubungkan nilai tengah tiap sisi atas yang berdekatan dengan nilai tengah jarak frekuensi mutlak masing-masing. Atau segi banyak (bidang rata yang sudut atau sisinya lebih dari empat)

*p-value* adalah suatu nilai probabilitas dalam menemukan uji statistik lebih besar dari yang dihasilkan

## R

**Random** adalah secara serampangan pengambilan sampling atau secara acak dari kelompok yang mewakili keseluruhan dari populasi yang lebih besar.

## S

**Sampel** adalah sebagian dari populasi yang, apabila diambil dengan benar, merupakan representasi dari populasi

**Sensus** adalah pengumpulan data pada seluruh populasi

**Statistik** adalah sebagai alat bantu untuk member gambaran atas suatu peristiwa melalui bentuk yang sederhana dapat berupa angka-angka atau berupa grafik-grafik.

**Statistika** adalah pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan atau penganalisisannya dan penarikan kesimpulan berdasarkan kumpulan data dan penganalisisan yang dilakukan

**Statistik Deskriptif** adalah sebagai alat bantu untuk mendiskripsikan fenomena-fenomena yang diteliti berdasarkan data yang terkumpul.

**Statistik Inferensial** adalah sebagai alat bantu tidak hanya untuk mendiskripsikan, tetapi lebih ditekankan pada fungsi analisis untuk menginferensialkan (menemukan ciri-ciri statistik tertentu) untuk suatu populasi dari suatu sampel secara random, dalam rangka pengujian hipotesis penelitian.

**Statistik Induktif** adalah statistik yang digunakan untuk membuat berbagai inferensi terhadap sekumpulan data yang berasal dari sampel.

**Statistik Deduktif** adalah statistic yang digunakan untuk membuat berbagai informasi terhadap sekumpulan data yang berasal dari berbagai sampel.

**Statistik Parametrik** adalah alat bantu analisis data dengan berdasar atas asumsi-asumsi, bahwa samplnya harus berdistribusi normal yang diambil secara random, dan datanya bersekala interval atau rasio.

**Statistik Nonparametrik** adalah alat bantu analisis data yang tidak harus memenuhi persyaratan-persyaratan seperti statistic parametric.

**Sampling** adalah proses pemilihan sejumlah individu atau serangkaian obyek pengukuran dari kelompok atau populasi yang adapemilihan sample yang berlain-lainan.

**Signifikasi** adalah pengertian atau mengandung arti penting.

**Standar Deviasi** adalah akar pangkat dua dari variansi yang merupakan bentuk linier yang nilainya selalu positif atau bisa juga disebut simpangan baku.

## T

**Teori** adalah pendapat yang dikemukakan sebagai keterangan mengenai suatu peristiwa (kejadian dsb)

## V

**Variabel** adalah Segala sesuatu yang berbentuk apa saja yang ditetapkan oleh peneliti untuk dipelajari sehingga diperoleh informasi tentang hal tersebut, kemudian ditarik kesimpulanya dapat berubah-ubah, berbeda-beda, bermacam-macam (tentang mutu, harga dsb).

**Varian** adalah salah satu teknik statistik yang digunakan untuk menjelaskan homogenitas kelompok.

**Verifikasi** adalah pemeriksaan tentang kebenaran laporan, perhitungan uang.

## BIOGRAFI PENULIS



**Azwar Habibi, S.Si., M.Si.**, lahir di Jember Jawa Timur pada tanggal 24 April 1985. Pada Bulan Januari 2021 diterima sebagai Dosen di IAIN Madura.

### **Riwayat Pendidikan Formal :**

- ✦ Pendidikan Sekolah Dasar di MIMA Salafiah-Syafi'iah XI Ambulu lulus tahun 1997.
- ✦ Pendidikan Sekolah Lanjutan Pertama di SLTP Negeri 1 Ambulu lulus tahun 2000.
- ✦ Pendidikan Sekolah Lanjutan Atas di SMU Negeri 1 Ambulu lulus tahun 2003.
- ✦ Pendidikan Strata-1 di Jurusan Matematika Fakultas Matematika dan Ilmu pengetahuan Alam Universitas Jember lulus tahun 2007.
- ✦ Pendidikan Strata-2 di jurusan Statistika Fakultas Matematika dan Ilmu pengetahuan Alam Institut Teknologi Sepuluh Nopember lulus tahun 2010