

**STATISTIKA DASAR:
KONSEPTUALISASI DAN APLIKASI**

Sanksi Pelanggaran Pasal 113

Undang-undang Nomor 28 Tahun 2014

Perubahan atas Undang-undang Nomor 7 Tahun 1987

Perubahan atas Undang-undang Nomor 6 Tahun 1982

Perubahan atas Undang-undang Nomor 19 Tahun 2002

Tentang Hak Cipta

- (1) Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000 (seratus juta rupiah).
- (2) Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).
- (3) Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp1.000.000.000,00 (satu miliar rupiah).
- (4) Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp4.000.000.000,00 (empat miliar rupiah). Pasal 114 Setiap Orang yang mengelola tempat perdagangan dalam segala bentuknya yang dengan sengaja dan mengetahui membiarkan penjualan dan/atau penggandaan barang hasil pelanggaran Hak Cipta dan/atau Hak Terkait di tempat perdagangan yang dikelolanya sebagaimana dimaksud dalam Pasal 10, dipidana dengan pidana denda paling banyak Rp100.000.000,00 (seratus juta rupiah). Pasal 115 Setiap Orang yang tanpa persetujuan dari orang yang dipotret atau ahli warisnya melakukan Penggunaan Secara Komersial, Penggandaan, Pengumuman, Pendistribusian, atau Komunikasi atas Potret sebagaimana dimaksud dalam Pasal 12 untuk kepentingan reklame atau periklanan untuk Penggunaan Secara Komersial baik.

STATISTIKA DASAR: KONSEPTUALISASI DAN APLIKASI

**Penulis:
Dr. Nalom Siagian, M. M.**

Statistika Dasar: Konseptualisasi dan Aplikasi

Copyright© 2021

Dr. Nalom Siagian, M. M.

viii + 194 hlm.; 17 cm x 24 cm

ISBN: 978-623-378-086-5

Penulis: Dr. Nalom Siagian, M. M.

Perancang Sampul: Fanny Hendro Aryo Putro, S. Sos., M.I.Kom.

Penyunting Naskah: Topan Setiawan, S. Sos., M.I.Kom.

Penata Letak: Siti Aisyah Hasanudin

Redaksi:

Kultura Digital Media

CV Kultura Digital Media

Jl Halilantar no 11, Ketingan RT 01 RW 11, Jebres, Surakarta

Pos-el: redaksikdm@gmail.com

Website: www.kulturadigitalmedia.com

Tlp. 081326279377 / 08112569111

Cetakan pertama: 2021

Perpustakaan Nasional RI Data Katalog dalam Terbitan (KDT)

Statistika Dasar: Konseptualisasi dan Aplikasi

penulis naskah, Dr. Nalom Siagian, M.M. - Surakarta: KDM, 2021.

viii + 194 hlm. ; 17 cm x 24 cm

ISBN: 978-623-378-086-5

1. Non-Fiksi: Statistika Dasar: Konseptualisasi dan Aplikasi. I. Judul
- II. Setiawan, Topan.

Hak cipta dilindungi oleh Undang-Undang

Dilarang mengutip atau memperbanyak sebagian

atau seluruh isi buku ini tanpa izin tertulis dari Penerbit

Isi di luar tanggung jawab Penerbit CV Kultura Digital Media

KATA PENGANTAR

Dengan segala kerendahan hati, tulus dan ikhlas, penulis memanjatkan Puji dan Syukur Kehadirat Tuhan Yang Maha Kuasa atas bimbingan Nya hingga penulisan buku yang berjudul Statistika Dasar : Konseptualisasi dan Aplikasi ini dapat terselesaikan.

Dalam menyelesaikan buku ini tentu saja penulis menemui banyak kendala dan hambatan, tapi berkat dorongan dan masukan dari berbagai pihak, buku ini dapat terselesaikan dengan baik. Untuk itu, penulis mengucapkan banyak terimakasih. Secara khusus kepada Istri tercinta Dra. Elida Napitupulu dan ananda Theresia Mastiur Ningsing Siagian, SPd, Grace Angelia Julieta Siagian yang masih kuliah dan Pangeran Goklas Jeremia Siagian yang selalu memberikan dorongan dan Doa.

Buku ini bertujuan untuk membantu mahasiswa, peneliti dan pekerja dalam memahami pengetahuan tentang statistika, membahas secara praktis berbagai persoalan penelitian kuantitatif melalui SPSS. Untuk memudahkan pembaca memahami konsep, buku ini dilengkapi dengan latihan soal yang akan membantu dalam pengolahan data.

Penulis menyadari, penerbitan buku ini masih terdapat beberapa kekurangan. Untuk itu, penulis menyampaikan terimakasih kepada para pembaca buku ini yang berkenan memberikan masukan demi kebaikan dalam penerbitan edisi berikutnya.

Medan, April 2021

Penyusun,

Dr. Nalom Siagian, M.M.

DAFTAR ISI

KATA PENGANTAR.....	ii
DAFTAR ISI	iii

BAB I

STATISTIKA: MAKNA, JENIS DAN PERANNYA DALAM PENELITIAN..... 1

1. Lingkup Statistika	1
1.1 Batasan Statistika	1
1.2 Perbedaan Statistik dan Statistika	2
1.3 Peran Elementer Statistika	5
1.4 Ekonomi dan Statistika	7
1.5 Statistika Dan Riset	8
2. Jenis Statistika	10
2.1 Statistik deskriptif	10
2.2 Statistik Inferensial.....	11
3. Penerapan/penggunaan statistika.....	12
3.1 Statistik terapan, statistik teoritis dan statistik matematika	12
3.2 Statistik dalam Sosial Kemasyarakatan	13
3.3 Komputasi Statistika.....	13
3.4 Statistik Sebagai Seni	13
4. Pembedaan Statistika Menurut Parameter	14
4.1 Statistika Parametrik.....	14
4.2 Statistika Nonparametrik	14
5. Statistika Dan Data	15
5.1 Data Kuantitatif vs Kualitatif	16

5.2	Data Menurut Dimensi Waktu	18
5.3	Data Menurut Cara Memperolehnya	19
5.4	Penyajian Data	22
BAB II		
DISTRIBUSI FREKUENSI		29
2.1	Batasan Distribusi Frekuensi	29
2.2	Membuat Tabel Distribusi Frekuensi	29
2.3	Distribusi Frekuensi Relatif dan Kumulatif	31
2.4	Histogram Frekuensi	35
2.5	Poligon frekuensi	36
BAB III		
KONSTRUKSI VARIABEL PENELITIAN DAN TEKNIK		
PENGUKURAN.....		38
3.1	Memahami Variabel Penelitian	38
3.2	Karakteristik Variabel Penelitian	39
3.3	Relasi Antar Variabel.....	40
3.4	Pengukuran Variabel Penelitian	42
3.5	Teknik Skala dan Pengumpulan Data	44
BAB IV		
POPULASI, SAMPEL, DAN TEKNIK PENGAMBILAN SAMPEL .		53
4.1	Populasi Dalam Riset	53
4.2	Kriteria & Proses Penyusunan Kerangka Sampel.....	54
4.3	Menentukan Ukuran Besaran Sampel.....	58

BAB V

PENTAHAPAN PENGUJIAN DENGAN STATISTIKA DALAM RISET..... 62

- 5.1 Pendugaan Dalam Penelitian 62
- 5.2. Pendugaan/Hipotesis Nihil, Hipotesis Kerja dan Pemilihan Uji Statistik 63
- 5.3 Menentukan Taraf Signifikansi dan Besaran Sampel Penelitian ... 64
- 5.4 Distribusi Sampling Dan Nilai Uji Statistik 65
- 5.5 Proses Pengambilan Keputusan Berdasar Hasil Pengujian 67

BAB VI

PENDUGAAN NILAI POPULASI DALAM STATISTIKA 71

- 6.1 Pendugaan Harga Rata-Rata Pupulasi Untuk Sampel $n \geq 30$ dan $n < 30$ 71
- 6.2 Pendugaan Proporsi Pada Populasi 74
- 6.3 Pendugaan Harga Beda *Dua Mean* Populasi dengan Sampel Besar dan Kecil 76

BAB VII

ANALISIS KORELASI 79

- 7.1 Konstruksi Analisis Korelasi Linear Sederhana 80
 - a. Data Nominal..... 80
 - b. Data Ordinal 80
 - c. Data Kontinum..... 81
 - d. Data Deret Waktu (*Time Series*)..... 82
 - e. Analisis Korelasi Linear Parsial 85
 - f. Analisis Korelasi Berganda (Multiple) 85

BAB VIII

ANALISIS REGRESI LINEAR.....	89
8.1 Analisis Regresi Linear Sederhana.....	89
8.2 Analisis Regresi Linear Berganda.....	96
8.3 Analisis Linear Berganda dengan Variabel <i>Dummy</i>	105

BAB IX

PROSEDUR PENGUJIAN STATISTIK NONPARAMETRIK.....	108
9.1 Statistik Nonparametrik.....	108
9.2 Beberapa Jenis Uji Statistik Nonparametrik.....	110
9.3 Uji Binomial.....	112
9.4 Uji tentang Proporsi Populasi.....	116
9.5 Uji Jenjang Bertanda Wilcoxon.....	117
9.6 Uji Walsh.....	124
9.7 Uji <i>Chi-Square</i> - Dua Sampel Independen.....	127
9.8 Uji Maan-Whitney.....	134
9.9 Uji Kolmogorov-Smirnov.....	143
9.10 Uji Friedman.....	148
9.11 Uji <i>Chi-Square</i> k Sampel Independen ($k > 2$).....	152
9.12 Uji Kruskal - Wallis.....	155
9.13 Koefisien Kontingensi C.....	160
9.14 Korelasi Rank Spearman.....	164

BAB X

KOMPUTASI DATA DENGAN PROGRAM SPSS.....	170
10.1 Sekilas Program SPSS.....	170
10.2 Mengoperasikan Program SPSS.....	171

10.3	Mengenal Windows SPSS	171
10.4	Input (Memasukkan) Data	175
10.5	Analisis Faktor dengan SPSS for Window	175
10.6	Analisis Korelasi dengan SPSS for Windows	177
10.7	Analisis Regresi dengan SPSS for Windows	179
10.8	Penyajian Hasil Analisis dan Interpretasi	180
DAFTAR PUSTAKA		183
DAFTAR TABEL.....		184
BIODATA PENULIS.....		194

BAB I

STATISTIKA: MAKNA, JENIS DAN PERANNYA DALAM PENELITIAN

1. Lingkup Statistika

1.1 Batasan Statistika

Terminologi statistika sebenarnya tidak asing lagi dalam diskursus masyarakat, terutama masyarakat akademik dan pihak-pihak yang bersentuhan dengan data, angka dan kuantitas. Dalam sudut pandang sejarah kata, terminologi statistika bersumber dari bahasa Latin, yakni “status” yang maknanya adalah negara. Dalam makna yang lebih luas lagi statistika adalah terminologi yang menyatakan hal hal yang berkaitan dengan konteks ketatanegaraan.

Jika di telusuri lebih jauh pada mulanya kajian statistika hanya berhubungan dengan kelompok kelompok simbol numerik atau angka tentang pemasukan/pendapatan masyarakat, tentang jumlah penduduk suatu wilayah. Dalam konteks ini menyangkut pula kelompok angka yang diperlukan oleh pemerintah dalam mencari exit strategi bagi masalah yang ada. Berbarengan dengan dinamika kehidupan yang kian kompleks, maka cakupan statistika juga semakin kompleks. Statistika merambah pada berbagai lini kehidupan praktis dengan tugas yang lebih bersifat pragmatis. Sementara kaitannya dengan kajian keilmuan, maka statistika dapat di katakan multidisiplin karena dapat diaplikasikan pada berbagai disiplin ilmu.

Problem perihal definisi atau batasan, umumnya tidak bisa langsung bersifat general, sebab sangat sulit untuk membawa kita pada penjelasan dan pemahaman yang komprehensif. Menurut Kamus Besar Bahasa Indonesia (KBBI) statistika yaitu ilmu yang mempelajari tentang cara mengumpulkan, menabulasi, menganalisis dan menemukan keterangan dari data. Biasanya, data yang dikumpulkan dalam statistika berupa angka-angka. Data – data tersebut pada akhirnya akan mengatarkan untuk sebuah proses pengambilan kesimpulan secara umum. Untuk menambah sudut pandang mengenai statistika, berikut adalah pemikiran Prof. Horace Secrist:

“Statistics is the aggregate of facts affected to a marked extent by the multiplicity of causes, numerically expressed, enumerated or estimated according to reasonable standards of accuracy, collected in a systematic manner for a pre-determined purpose and placed in relation to each other (Deilber, 1943).”

Batasan yang di kemukakan oleh pakar matematika, Prof. Horace Secrist, memuat pengertian bahwa statistik adalah kumpulan fakta yang dipengaruhi sampai batas tertentu oleh banyaknya penyebab, dinyatakan secara numerik, dihitung atau diperkirakan menurut standar akurasi yang wajar, dikumpulkan secara sistematis untuk tujuan yang telah ditentukan sebelumnya dan ditempatkan dalam kaitannya satu sama lain.

Dalam batasan itu yang paling menonjol adalah akurasi dan sistematika didalam prosesnya. Akurasi yang wajar menjadi hal yang elementer sebab pada akhirnya akan bertaut dengan derajat kebenaran faktual suatu realitas yang akan diukur. Sistematika juga menjadi hal yang wajib di perhatikan agar sajian data yang dihasilkan terlihat rapi dan logis, dengan herarki pemikiran terstruktur.

1.2 Perbedaan Statistik dan Statistika

Pada mulanya cenderung akan tak disadari sebagai hal yang terlihat sama, betapa sesungguhnya terdapat perbedaan statistik dan statistika yang perlu diketahui. Kedua kata tersebut banyak digunakan pada saat menulis laporan hasil kajian atau riset tertentu. Dalam proses kerja pengolahan data yang dikumpulkan secara numerik-pun, statistik dan statistika juga sering dijumpai pada saat membuat hasil dari pengolahan data tersebut. Pendek kata statistik dan statistika sangat dibutuhkan dalam menyelesaikan masalah – masalah tertentu. Akan tetapi keduanya memiliki cakupan makna yang berbeda, hingga masih ada orang yang cenderung keliru saat penyebutan dan penempatan maknanya dalam sebuah wacana tertentu.

Untuk memahami lebih jauh mengenai perbedaan kedua istilah tersebut dapat ditentukan dengan parameter antara lain dari batasan/definisi istilah, tujuan dan metode yang digunakan. Berikut ini adalah perbedaan kedua istilah tersebut dilihat dari tiga sudut pandang yakni batasan, tujuan dan metodenya.

1.2.1 Pengertian Statistik dan Statistika

Dalam wacana ini, perbedaan kedua istilah itu dilihat dari pengertian kedua istilah tersebut secara teoretis. Terminologi statistik pada dasarnya merupakan hasil data yang di visualisasikan / ditampilkan dalam berbagai ragam bentuk tabel dan grafik. Menurut Kamus Besar Bahasa Indonesia (KBBI) statistik adalah angka-angka atau catatan yang dikelompokkan, dikumpulkan serta ditabulasi, sehingga akan didapatkan informasi yang berkaitan dengan masalah tertentu.

Realitas numerik atau data pada statistik bisa dihimpun dari proses pengumpulan data melalui wawancara atau data skunder lainnya. Statistik pada dasarnya merupakan kumpulan data yang dihasilkan dari pengolahan yang dipelajari dalam statistika. Dengan kata lain, statistika merupakan prespektif ilmunya, sementara statistik adalah hasil olahan datanya.

Selanjutnya, mengenai batasan atau pengertian statistika dapat dipahami sebagai metode atau sebuah bidang keilmuan yang mempelajari mengenai bagaimana menghimpun, mengumpulkan, membuat rancangan, merencanakan, menginterpretasi dan menganalisa data empiris dari lapangan. Tujuannya untuk menampilkan atau mempresentasikan hasil atau data yang diperoleh dari realitas empiris lapangan. Singkatnya, statistika merupakan salah satu cabang ilmu pengetahuan yang terkait dengan data-data, cara mengoleksinya serta bagaimana menampilkannya sehingga mudah dipahami oleh orang.

1.2.2 Tujuan Statistik dan Statistika

Menentukan perbedaan statistik dan statistika yang kedua akan dilihat dengan sudut pandang tujuannya. Kedua istilah itu sama sama bertautan dengan proses ilmiah pengumpulan, pengolahan serta interpretasi data, namun dapat dirasakan betapa sebenarnya tujuan statistik dan statistika itu berbeda. Sebagaimana diurai diatas, statistik merupakan hasil data yang ditampilkan dalam bentuk tabel hingga grafik. Data mentah yang telah diolah serta ditampilkan inilah yang kemudian disebut dengan statistik. Tujuan dan fungsi statistik itu sendiri adalah guna mendapatkan deskripsi atau gambaran atas data – data yang telah dikumpulkan, dihimpun dan dikaji sebelumnya. Dengan begitu, dari sajian data tersebut dapat dibuat simpulan atau

generalisasi, mengenai problem atau fokus permasalahan yang sedang di teliti.

Pada sisi yang lain, statistika merupakan sebuah metode ilmiah yang digunakan untuk mempermudah dalam pengolahan dan menginterpretasi data yang sudah dikumpulkan. Dengan menggunakan ilmu statistika, data mentah empiris dilapangan sebuah masalah bisa diolah, diinterpretasikan dan digunakan untuk tujuan tertentu. Data hasil dari olah statistika ini, nantinya akan disebut secara luas sebagai data statistik. Maka itu, jika di cermati fungsi statistika secara umum yakni untuk mengonversi atau mengubah data serta informasi yang acak sifatnya (Row Material) menjadi sebuah sajian data statistik yang bisa dipahami.

1.2.3 Metode yang Digunakan

Dimensi ketiga pembedaan kedua istilah tersebut adalah melalui metode yang digunakan ketika dilakukan proses kerja. Statistika dapat menggunakan metode penelitian survei dan eksperimen pada saat proses pengumpulan data. Sedangkan statistik hanya menangkap hasil pengolahan data yang dilakukan oleh statistika dalam bentuk angka jadi yang dituangkan misalnya sebagai grafik, diagram dan tabel. Data yang telah tersaji tersebut dinamakan data statistik. Jadi sebenarnya statistik tidak secara langsung menggunakan metode ilmiah.

Setelah memahami perbedaan statistik dan statistika, dan agar lebih jelas mendapatkan gambaran perbedaan, berikut contoh – contohnya dalam kehidupan nyata. Contoh statistik pada kehidupan sehari-hari: (1) Data kepemilikan kendaraan bermotor di wilayah tertentu, (2) Data kependudukan dan perekonomian milik Badan Pusat Statistik (BPS), (3) Data kependudukan suatu desa melalui survey, (4) Data jumlah sekolah di suatu propinsi dan lain sebagainya.

Sementara itu contoh statistika pada kehidupan sehari-hari: (1) Pengolahan data kependudukan kawasan tertentu untuk menentukan transisi demografi, (2) Pengolahan data kependudukan wilayah tertentu, (3) Menemukan rata-rata serta standar deviasi atas nilai ujian mahasiswa dan sebagainya.

1.3 Peran Elementer Statistika

Eksistensi statistik dalam dunia riset utamanya riset-riset dalam perspektif kuantitatif, pasti menggunakan statistika sebagai alat bantu. Hal ini menjadi niscaya sebab kajian kuantitatif memerlukan perhitungan yang harus ditopang oleh statistika guna menganalisis secara numerik problem yang hadir. Dengan demikian baik secara konsep maupun teorinya, statistika menjadi menu yang wajib ada. Peruntukan itu misalnya adalah guna mengevaluasi apakah cara, rumus, metode atau hal-hal yang baru ditemukan menampilkan performa yang lebih baik daripada cara lama, melalui riset yang dilakukan dilaboratorium atau penelitian yang dilakukan dilapangan. Dapat juga untuk melihat apakah konstruksi, model, rumusan dan pola untuk sesuatu bidang atau hal tertentu dapat kita pakai dan anut atau sudah tidak lagi berlaku.

Secara teoretis, statistika juga memiliki kemampuan prediktif, sehingga dalam konteks ini dapat digunakan untuk melihat realitas masa depan, tentu saja dalam sudut pandang prediktif dengan tetap memiliki derajat keliru tertentu. Statistika juga sanggup meramalkan apakah sebuah faktor yang satu dipengaruhi atau mempengaruhi faktor-faktor yang lainnya. Andaikata terdapat hubungan antara faktor-faktor tersebut, berapa kuat derajat atau tingkat adanya hubungan tersebut. Statistika juga dapat untuk menguji apakah hipotesis penelitian yang diajukan akan dapat diterima atau malah ditolak secara ilmiah.

Sekadar menengok sejarah penggunaan statistika, posisi dan arti statistika sebagai alat bantu riset dalam menjelaskan permasalahan penelitian sangat strategis dan elementer, terutama semenjak kepoloporan Karl Pearson. Dalam masa itu, sekitar abad ke 19 awal, Pearson memantik lalu memperkenalkan bagaimana metode statistika diterapkan melalui konsep regresi, korelasi, distribusi Chi-kuadrat dan analisis statistika untuk menolah data penelitian.

Lebih dari itu perspektif metodis dari statistika modern telah dengan meyakinkan dipergunakan di bidang biologi, pertanian, dan ekonomi merupakan produk abad ke dua puluh dan memperoleh kemajuan yang sejak tahun 1918-1935 ketika R. Fisher memperkenalkan analisis varian kedalam literatur statistika (Dajan, 2000). Semenjak tonggak itulah penggunaan metode statistika makin meluas dari bidang ilmu hayat atau biologi dan

agraris atau pertanian ke bidang-bidang keilmuan lainnya. Bidang-bidang ilmu pengetahuan yang lain seperti ranah biometri, kajian agronometri, ilmu ekonometrik, pengukuran dalam psikometri, sosiometri, dan bahkan hingga ilmu antropologi yang kemudian disebut antropometri telah semakin “nyaman” dengan pendekatan statistika.

Riset yang berkualitas mungkin tidak sekali hadir, ia pasti memerlukan perencanaan yang baik di awal dan *on-going evaluation* yang konsisten. Dalam konteks ini perencanaan dan evaluasi hasil kajian/penelitian secara statistika memberikan ruang leluasa bagi ilmuwan untuk melakukan revisi, koreksi, perbaikan dan penyempurnaan atas hasil kajian atau prototipe yang dihasilkan bagi penggunaan yang semakin sempurna.

Fungsi kontrol numerik yang dilakukan peneliti dengan bantuan statistika, akan menjamin hasil penemuan yang berguna bagi kepentingan umat manusia semakin berkualitas. Kontrol kualitas dalam sudut pandang statistika ini dalam berbagai perusahaan jasa maupun barang bukan saja dapat mempertahankan kualitas produknya, tetapi memberikan kontribusi pada kemungkinan perbaikan kualitas barang-barang produksi itu sendiri. Pada bidang lain kemanfaatan elementer ilmu statistika adalah dalam bidang kesehatan, kedokteran, epidemik dan sebagainya. Pemodelan statistik dapat memberikan data data penting yang dapat digunakan untuk mengambil keputusan.

Jika di bentang menurut ukuran waktu, setidaknya dalam interval waktu setengah abad terakhir ilmu statistika mengalami perkembangan yang sangat signifikan. Kajian – kajian saintifik yang berbasis survey, laboratoris hingga eksperimental akan terbantu secara lebih teliti dengan penggunaan metode statistik yang disesuaikan dengan kebutuhan yang muncul.

Peran elementer statistika yang perlu kita sadari bersama sesungguhnya adalah dalam bidang pendidikan, dalam konteks ini utamanya adalah pendidikan dalam bidang studi yang secara langsung berkaitan dengan hitung menghitung. Realitas atau fakta menunjukkan betapa kurangnya fokus perhatian terhadap tingkat kemampuan berpikir secara kreatif dalam mata pelajaran matematika beserta implikasinya benar-benar terjadi dalam diskursus pendidikan di Indonesia. Realitas itu kemudian menuntut adanya upaya untuk memberikan perhatian lebih pada tingkat kemampuan kreatif tersebut dalam pembelajaran matematika dewasa ini. Langkah maju dalam

pengembangan tingkat kemampuan berpikir mata pelajaran statistik merupakan salah satu tema sentral dari upaya pembelajaran matematika yang hebat lagi, sehingga diharapkan mampu untuk memantik dan menumbuhkan daya kreativitas siswa dalam pembelajaran bidang ilmu matematika.

Dalam perspektif Suriasumantri (2003), berpikir secara statistik pada dasarnya merupakan bagian dari sarana berpikir ilmiah, yang didalamnya terdapat bagian dari berpikir logis dan sistematis. Didalam proses belajar statistik tingkat penguasaan kompetensi sangatlah penting, karena dapat dirasakan betapa belajar statistik menjadi prasyarat utama bagi para siswa untuk mengetahui kemampuan berpikir statistiknya. Dengan menguasai konsep berpikir statistik maka dapat diharapkan akan membantu siswa dalam memahami pelajaran matematika yang selanjutnya siswa diharapkan memiliki kemampuan berpikir logis, analitis, sistematis, kritis, dan kreatif.

Kemampuan berpikir statistik sebenarnya berada pada seluruh area matematika dan begitu penting dalam aktivitas kehidupan sehari-hari. Pemikir, seperti H.G. Wells (Suriasumantri, 2003) mengatakan betapa kelak, suatu hari nanti berpikir statistik akan menjadi keharusan bagi manusia seperti juga membaca dan menulis. Pemikiran ini jelas bukan pemikira yang berlebihan, sebab dengannya siswa diproyeksikan dapat menggunakan daya nalar yang logis, analitis, sistematis, dan kreatif. Lebih dari itu diharapkan siswa mampu menemukan *exit strategy* sebagaimana yang dibutuhkan dalam penyelesaian permasalahan.

1.4 Ekonomi dan Statistika

Pada praktik kehidupan, dunia ekonomi sangat berkaitan dengan ilmu statistika. Manusia sebagai anima ekonomika, adalah entitas yang tidak lepas dari aktivitas perekonomian dengan berbagai bentuknya. Dalam melakukan aktivitas perekonomian, manusia memerlukan alat bantu untuk mempermudah perhitungan dan membuat analisis bisnis yang tepat agar proses bisnis dan perekonomian sesuai dengan tujuannya. Salah satu alat bantu yang representatif dalam konteks ini adalah ilmu statistika.

Penggunaan statistika dalam bidang ekonomi, misalnya dalam perusahaan adalah salah satu patokan dalam mengambil keputusan bisnis yang penting. Keputusan bisnis penting dan tepat akan berguna untuk

memastikan jalannya perusahaan lebih terkendali secara efektif dan efisien. Sekadar ilustrasi, dalam sebuah perusahaan, terdapat divisi pemasaran baik barang dan jasa yang wajib mengetahui segmentasi pasar yang akan dituju. Dalam hal ini sudah pasti dibutuhkan data statistik, mengenai profil pasar yang ditargetkan. Pada titik inilah data statistik diperlukan untuk dunia pemasaran, sehingga ilmu statistika tampak memberi peran yang berarti.

Untuk menentukan kebijakan pemasaran produk tertentu sesuai dengan tepat data-data konsumen seperti jumlah penduduk suatu wilayah/daerah, kriteria masyarakat berdasar usia atau jenis kelamin, pekerjaan beserta penghasilan masyarakat, dan sebagainya sangat diperlukan. Ilustrasi lain adalah misalnya kebijakan mengenai kenaikan gaji atau tunjangan lainnya misalnya, maka diperlukan profil statistik berkaitan dengan segala pembiayaan operasional perusahaan tersebut. Dari profiling tersebut akan di dipetakan peluang besaran dana yang akan dialokasikan untuk kenaikan gaji karyawan. Lebih detil lagi adalah akan dibutuhkan data – data yang secara langsung berhubungan dengan profil karyawan/pegawai yang ada. Tentu akan dilakukan klasifikasi pegawai berdasarkan misalnya adalah masa kerja, golongan/pangkat, jabatan, derajat pendidikan. Profil karyawan berdasar data itu akan di sajikan dalam bahasa statistik seperti diagram, tabel dan sebagainya. Dengan demikian pimpinan pengambil kebijakan dapat menentukan keputusannya berdasarkan data-data tersebut, sehingga diperoleh keputusan yang diharapkan berkeadilan.

1.5 Statistika Dan Riset

Tradisi penelitian secara garis besar ada dua, yakni kualitatif dan kuantitatif. Tradisi tersebut berbeda , terutama dalam diskursus ini adalah bila dilihat dari penggunaan ilmu statistika. Penelitian kualitatif relatif tidak memerlukan perhitungan numerik dan proses lainnya sebagaimana diperankan oleh ilmu statistika. Kalaupun ada penggunaan umumnya tidak dominan. Berbeda dengan penelitian positivistik yang berjalan dalam koridor metode kuantitatif, maka penggunaan ilmu statistika dalam riset ini begitu dominan. Dalam riset dengan paradigma positivistik, maka penelitian kuantitatif ini mencoba meneropong suatu gejala itu sebagai sesuatu yang nyata jika bisa dilihat, diukur, dan diklasifikasikan. Dengan demikian,

diskursus yang dibangun dalam penelitian ini adalah diskursus untuk membuat gamblang peran dominan penelitian kuantitatif.

Proses penelitian, dalam konteks ini riset kuantitatif merupakan penelitian yang dilakukan dengan cara ilmiah. Cara ilmiah dalam melaksanakan penelitian atau riset sesungguhnya meniscayakan langkah-langkah kerja yang sifatnya sistematis. Langkah kerja yang sistematis ini mempunyai profil terstruktur yang ketat. Hal ini untuk memberikan jaminan metodologis, sehingga dapat dipertanggungjawabkan secara ilmiah. Proses penelitian akan dilakukan pertama kali dengan upaya mengidentifikasi permasalahan yang muncul. Masalah pada dasarnya merupakan gap atau perbedaan antara *das sollen* atau *realitas ideal* yang diharapkan, dengan *das sein* atau realitas faktual apa yang terjadi. Masalah ditampilkan dalam *problem statement* yang spesifik, sehingga menjadi lebih fokus ketika peneliti mencari jawaban melalui pelaksanaan riset tersebut.

Selanjutnya, masalah inilah yang kelak dicari jawabannya yang dipandu dengan konstruksi teoretik yang kompatibel dan memadai. Kerangka teoretik ini tidak boleh dirumuskan secara sembarangan, karena teori merupakan pisau analisis yang sangat diperlukan dalam membahas permasalahan yang telah dipancarkan di awal penelitian. Pada lazimnya penulisan kerangka teoretik berdasarkan temuan-temuan atau teori-teori yang sudah ada pada penelitian – penelitian sebelumnya.

Dalam praktiknya hipotesis dapat juga dianggap sebagai teori meskipun masih dalam kategori sangat lemah. Mengapa sangat lemah karena hipotesis masih memerlukan pembuktian, yakni dengan pengujian dalam penelitian tersebut. Hipotesis pada dasarnya merupakan jawaban yang sifatnya masih sementara atas pertanyaan riset sebelumnya telah dirumuskan. Selanjutnya proses penelitian dilakukan dengan mengumpulkan fakta-fakta di lapangan untuk dapat membuktikan apakah teori sementara (yang dalam konteks ini adalah berupa hipotesis atau dugaan sementara) yang telah dirumuskan tersebut terbukti benar atau salah.

Untuk memberi gambaran yang lebih jelas mengenai peran statistik dalam riset berikut akan dirunut perannya sesuai dengan tahapan sebuah penelitian. Secara umum penelitian melibatkan statistika menurut 4 tahapan yang berurutan. Selanjutnya dapat di proyeksikan bahwa sejalan dengan proses penelitian, maka 4 tahapan dapat dijelaskan dalam hal ini adalah peran-

peran yang dilakukan oleh ilmu statistika sebagai berikut: (1) Penentuan sampel. Statistik diperlukan untuk menghitung berapa jumlah sampel yang diperlukan agar merepresentasikan populasi, (2) Mengumpulkan data. Dalam pengumpulan data diperlukan suatu instrumen yang valid dan reliabel. Statistik berperan dalam menguji apakah instrumen yang digunakan sudah valid dan reliabel, (3) Menyajikan data. Dalam menyajikan data yang begitu banyak, diperlukan suatu teknik agar data tersebut lebih interaktif dan mudah dipahami oleh orang awam, misalnya dengan dibikin tabel atau diagram, dan (4) Analisis data. Tentu saja kegunaan utama statistik adalah untuk analisis data dalam menguji hipotesis. Dalam hal ini ada beberapa teknik yang digunakan, seperti regresi, anova, SEM, dll.

2. Jenis Statistika

Pada dasarnya penggunaan ilmu statistika sebagai alat panduan dasar, titik acuan dan wahana pengumpulan informasi, biasanya menggunakan ilmu statistika sebagai metode yang disebut sebagai berikut:

2.1 Statistik deskriptif

Statistika deskriptif merupakan langkah awal dalam pembahasan statistika yang hanya berhubungan dengan hal menguraikan atau memberikan keterangan-keterangan mengenai suatu keadaan atau fenomena. Dengan kata lain, statistika deskriptif hanya berfungsi menerangkan keadaan, gejala, dan fenomena. Selanjutnya dapat dikatakan, penggunaan statistika deskriptif tidak dapat mengambil kesimpulan yang bersifat umum (generalisasi) karena statistika di sini memang terbatas pada hal yang ada saja. Hasil analisisnya masih sederhana, bahkan sebagian besar analisis atau perhitungannya bersifat penyederhanaan atas data yang terkumpul. Berdasarkan hasil penyederhanaan inilah pemakai diharapkan dapat mengambil arti dari data tersebut. Didasarkan atas ruang lingkup bahasannya, statistika deskriptif mencakup hal berikut: (1). Distribusi frekuensi beserta bagian bagiannya, seperti: grafik distribusi (histogram, poligon frekuensi, dan ogive) ; ukuran nilai sentral (rata-rata, modus); ukuran letak (median, kuartil, desil, dan persentil); ukuran dispersi (nilai jarak, simpangan rata-rata, variasi, simpangan baku, dan sebagainya), (2) Deret waktu atau data berkala, (3)

Kolerasi dan determinasi, (4) Regresi (sederhana dan berganda), (5) Angka Indeks.

Pada metode ini ilmu statistika merupakan wahana atau sarana untuk memberikan informasi tentang kinerja proses tertentu yang sedang berjalan. Dalam statistik deskriptif, benda yang sifatnya bisa dihitung disajikan dalam angka, tabel maupun diagram yang merupakan representasi dan ringkasan yang secara kuantitatif menggambarkan atau merangkum kumpulan informasi. Statistika deskriptif bertujuan untuk merangkum sampel, bukan menggunakan data – data yang ada untuk mempelajari mengenai populasi yang dianggap mewakili sampel data tersebut.

2.2 Statistik Inferensial

Statistika inferensial merupakan bagian dari ilmu statistika yang lebih fokus mempelajari mengenai interpretasi dan pengambilan kesimpulan (generalisasi) yang berlaku secara umum dari data-data yang telah dikumpulkan. Statistika inferensial bertalian dengan pendugaan besaran populasi dan pengujian hipotetis dari suatu realitas tertentu atau fenomena. Dengan kalimat lain, statistika inferensial berfungsi untuk memprediksi atau meramalkan dan mengontrol atau mengatur suatu kejadian.

Statistik inferensial pada dasarnya memberikan informasi tentang prediksi yang berhubungan dengan kinerja sebuah proses atau peluang. Inferensial adalah proses analisis data dengan menggunakan statistika untuk membantu menyimpulkan mengenai permasalahan yang di teliti. Analisis statistik inferensial mampu menyimpulkan sifat-sifat suatu populasi, misalnya dengan cara menguji hipotesis dan memperoleh taksiran mengenai data – data yang akan diukur. Dalam konteks statistika inferensial ini, kumpulan data yang akan diambil cuplikannya atau sampelnya, Diasumsikan bahwa kumpulan data yang diamati diambil sampelnya dari populasi yang lebih besar. Statistik inferensial dalam hal ini terlihat kontras dengan statistik deskriptif. Statistik deskriptif lebih kepada data yang semata-mata berhubungan dengan sifat sifat data yang teramati. Sementara statistika inferensial didasarkan pada asumsi bahwa data tersebut berasal dari populasi yang lebih besar.

Proses menyimpulkan pada statistika inferensial ini pada dasarnya merupakan proses pembuatan generalisasi dari suatu populasi berdasarkan

realitas faktual data (sampel) yang ada, yang dikoleksi dari lapangan. Berdasarkan lingkup pembahasannya, statistika inferensial meliputi: (1) Probabilitas atau teori peluang, (2) Distribusi teoretis, (3) Sampling dan distribusi sampling, (4) Pendugaan populasi atau teori populasi, (5) Uji hipotesis, (6) Analisis korelasi dan uji signifikansi, (7) Analisis regresi untuk peramalan.

Selanjutnya berdasarkan orientasi pembahasannya, maka terdapat perbedaan menjadi dua, yakni statistika matematis dan statistika terapan. Statistika matematika – sama dengan konsep statistika teoretis, yang lebih mengutamakan pemahaman akan model, derivasi konseptual, kumpulan rumus statistika dalam kerangka matematis-teoretis, pemahaman dan penggunaan uji-t, analisis regresi, uji normalitas dan homogenitas, galat dan lainnya. Selanjutnya adalah statistik terapan yang mengutamakan mengenai pemahaman atas konsep dan teknik teknik statistika yang aplikasinya hanya berlaku pada disiplin ilmu tertentu saja.

3. Penerapan/penggunaan statistika

Berbicara mengenai penerapan ilmu statistika, secara umum statistika cenderung digunakan oleh orang yang bekerja dalam bidang ilmu dan usaha dalam skala besar untuk mengukur banyak hal yang berhubungan dengan perhitungan dan visualisasi data. Para pengguna statistika bisa berada dalam lembaga swasta, seperti lembaga – lembaga survey, perusahaan swasta, maupun lembaga pemerintahan. Lembaga pemerintah yang jelas menggunakan ilmu statistika dalam menjalankan fungsi utamanya adalah Badan Pusat Statistik. Secara empiris, bukan hanya lembaga lembaga berskala besar saja yang dalam menjalankan kerjanya menggunakan ilmu statistik, namun ada juga lembaga yang sering menggunakan statistik dalam skala kecil dengan tujuan kecil pula, misalnya mahasiswa yang sedang menyusun tugas akhir dengan penelitian kuantitatif. Berikut ini adalah beberapa penerapan penggunaan statistika dalam kehidupan:

3.1 Statistik terapan, statistik teoritis dan statistika matematika

Dalam kehidupan sehari-hari, statistika dapat dibedakan dalam hal implementasinya. Statistik terapan sebagaimana dikatakan di atas adalah dua statistik yang berbentuk deskriptif dan inferensial sedangkan statistik teoritis

menyangkut argumen logis yang mendasari justifikasi pendekatan terhadap inferensi statistik, serta mencakup statistik matematika. Statistik matematika tidak hanya mencakup manipulasi distribusi probabilitas yang diperlukan untuk memperoleh hasil yang terkait dengan metode estimasi dan inferensi, tetapi juga berbagai aspek statistik komputasi dan desain eksperimen.

3.2 Statistik dalam Sosial Kemasyarakatan

Pada kenyataannya, ilmu statistika berlaku untuk berbagai bidang atau disiplin ilmu akademik, sehingga penggunaannya dapat dikatakan tersebar pada berbagai kalangan, seperti organisasi kemasyarakatan, organisasi bisnis/perusahaan. Pemerintahan dan sebagainya. Beberapa konsultan tentang statistika di butuhkan jasanya untuk membantu berbagai lembaga tersebut, manakala tidak memiliki tenaga ahli di kalangan internal mereka, khususnya menyangkut permasalahan spesifik yang dihadapi perusahaan/ organisasi tersebut.

3.3 Komputasi Statistika

Dunia teknik digital untuk keperluan analisis data dan pengukuran semakin canggih. Realitas peningkatan yang sangat cepat dan kontinu itu dimulai sekitar paruh abad 20 dan daya komputasi digitalpun semakin kuat berdampak pada praktik penggunaan ilmu statistika. Pada awal perkembangannya, pemodelan statistika lebih bersifat linear namun kini bertransformasi menuju model komputasi yang kuat dengan algoritma numerik yang kompatibel maka penggunaan ilmu statistika terbiasa da terpola juga pada model nonlinelar, seperti jaringan saraf buatan/*artificial neural networks*.

3.4 Statistik Sebagai Seni

Sudut pandang lain dalam melihat ilmu statistika adalah menggunakan perspektif seni. Dalam konteks ini ilmu statistika dinikmati sebagai seni mengolah data dan menggambar kesimpulan penelitian. Ilmu statistika diposisikan pada mulanya hanya sekadar sebagai metodologi yang sifatnya semi-standar yang diperlukan pada proses pembelajaran pada sebagian besar disiplin ilmu. Tradisi ini lantas berubah dengan mulai diperkenalkan penggunaan statistika dalam diskurusus atau konteks yang non-inferensial.

4. Perbedaan Statistika Menurut Parameter

Berdasarkan bentuk parameternya, statistika dapat dibedakan menjadi 2 (dua) bentuk, sebagai berikut :

4.1 Statistika Parametrik

Statistika parametrik adalah ilmu statistika yang mengharuskan persyaratan yang sifatnya khusus - spesifik mengenai parameter-parameter yang dibutuhkan, dan juga menghendaki populasi dimana sampel akan diambil. Dalam praktiknya, semua syarat ini tidak akan diuji, hanya diduga saja bahwa syarat tersebut dapat dipenuhi. Kemudian, hasil dari suatu pengujian parametrik akan dapat diterima berdasarkan suatu harga penerimaan (validitas) dari asumsi yang telah ditentukan sebelumnya. Pada sisi yang lain, di dalam uji parametrik dibutuhkan pula derajat pengukuran yang teliti.

4.2 Statistika Nonparametrik

Statistika non parametrik adalah jenis ilmu statistika dengan model atau corak yang tidak memerlukan parameter spesifik atau khusus dari populasi yang akan observasi atau diamati. Uji nonparametrik tidak membutuhkan pengukuran dengan derajat ketelitian presisi sebagaimana uji parametrik, sehingga uji nonparametrik biasa diterapkan untuk menganalisis data pada skala-skala ukuran ordinal dan nominal.

Dalam memilih salah satu uji statistika untuk menganalisis data harus diperhatikan kriteria sebagai:

1. Tingkat kemampuan pengujian.
2. Kemampuan uji tersebut dalam menganalisis data.
3. Efisiensi kemampuan uji.
4. Taraf pengukuran yang dapat dicapai dalam pengambilan data pengamatan.

Dalam khasanah ilmu-ilmu sosial cukup sulit upaya kita untuk menentukan suatu cara pengukuran tertentu, yang dapat diuji dengan uji parametrik, maka pada titik inilah uji nonparametrik akan digunakan. Dalam penerapan beberapa uji statistik nonparametrik, biasanya pencacahan ditulis dalam bentuk “rank” atau rangkaian.

5. Statistika Dan Data

Pada dasarnya rumus - rumus statistika dapat diterapkan jika rangkaian data yang tersaji adalah data dalam bentuk kuantitatif. Sebuah data dapat dikatakan sebagai data kuantitatif adalah bila serangkaian observasi atau pengukuran terhadap realitas dapat dinyatakan kedalam angka-angka tertentu. Dalam konteks ini contohnya adalah berat badan siswa sekolah yang dapat dinyatakan dengan 50 kg, 56 kg, 70 kg dan lain sebagainya. Jika diamati data kuantitatif tersebut masih bersifat acak, merupakan data asli (raw data) karena langsung dihimpun dari hasil pengukuran. Pada umumnya masih berbentuk catatan saja, dan belum mengalami pengolahan atau pengelolaan sehingga menjadi tertata.

Dalam kenyataannya tidak semua kegiatan pengumpulan data melalui observasi atau aktivitas pengukuran itu selalu bersifat kuantitatif, namun ada juga observasi atau pengamatan yang bersifat kualitatif merupakan serangkaian observasi yang terdapat dalam sampel (atau populasi), yang terhimpun dalam kelas-kelas yang saling lepas pisah. Dalam keadaan itu tidak mungkin untuk dinyatakan dalam angka – angka. Meskipun data yang dihimpun tersebut bersifat kualitatif bukan sama sekali tak bisa di di manfaatkan untuk melakukan analisis statistika. Pencacahan frekuensi jawaban yang berbeda dan persentasi kuantitas observasi yang tergolong dalam kelas yang berbeda tersebut sebenarnya sudah merupakan manifestasi analisis statistika. Pada dasarnya data yang bersifat kualitatif di atas dapat diklasifikasi ulang sebagai data kuantitatif. Data yang bentuknya angka-angka sebenarnya merupakan komponen elementer bagi setiap penelitian yang sifatnya statistika. Para ilmuwan statistika, pada umumnya akan menganggap betapa langkah mula – mula setiap riset/penelitian statistika ialah pengumpulan data dari realitas yang masih asli atau realitas natural. Pada titik ini diyakini bahwa baik dan buruknya, ketat tidaknya, cara pengumpulan data kasar atau *raw data* ini akan memengaruhi hasil dari pengolahan dan analisis data tersebut selanjutnya. Hal inilah yang dalam penelitian kuantitatif kecermatan dan kehati-hatian sudah dimulai sejak awal pengumpulan data kasar tersebut. Tujuannya adalah untuk mendapatkan asupan data yang benar-benar sesuai kondisi apa adanya dilapangan dan kompatibel dengan kajian atau riset yang dilakukan.

Secara definitif, data merupakan angka, fakta, fenomena, keadaan, atau realitas lainnya yang di konstruksi menurut logika tertentu, dan merupakan hasil dari proses pengamatan, pengukuran atau pencacahan atas variabel dari suatu objek kajian. Data berfungsi untuk membedakan objek yang satu dengan lainnya pada variabel yang sama. Jika kita berada dalam kajian ilmu ekonomi, dapat kita pahami bahwa data merupakan sekumpulan informasi yang digunakan untuk pembuatan keputusan dan pengambilan kebijakan tertentu. Dalam aktivitas riset, data perlu di konstruksi atau disusun dan disimpan dengan menggunakan metode tertentu. Hal ini agar kelak jika dibutuhkan dapat di cari dengan cepat dan tetap terstruktur.

5.1 Data Kuantitatif vs Kualitatif

Data – data dihipun dengan cara mengukur nilai variabel dalam sampel atau populasi yang telah ditentukan. Semua data tersebut selanjutnya dapat digolongkan menjadi dua jenis, yakni data kuantitatif dan data kualitatif. Data kuantitatif salah satu cirinya adalah data dapat yang diukur dalam suatu skala numerik (angka, bilangan). Data kuantitatif dapat dibedakan menjadi 2 jenis yakni sebagai berikut:

1. Data interval, data ini di definisikan sebagai sebuah skala yang mempunyai ciri-ciri skala ordinal. Ciri berikutnya adalah selisih dari setiap besaran angka dalam skala tersebut diketahui besarnya. Selanjutnya proses pengukurannya dilakukan dengan lebih rigid dan cermat. Dalam skala interval unit pengukuran bersifat tetap dan ditandai dengan oleh angka-angka sebagai representasi objek yang ditampilkan. Dalam proses pengukuran, perbandingan dari dua interval bukan bergantung pada unit pengukuran dan adanya titik asal, yakni titik nol. Artinya konstruksi data dalam skala interval tidak memiliki titik nol konstan atau pasti. Gambarannya adalah skor nilai dalam ulangan akhir semester berentang 0 hingga 100; ukuran panjang meteran berkisar 0 cm hingga 1000 cm dan lain sebagainya. Skala interval mempunyai sifat “unique to linear transformation” yang tidak terpengaruh jika dikalikan dengan suatu angka tetapan positif.
2. Data rasio, hal ini cirinya adalah bila suatu skala interval memiliki titik nol yang real. Di dalam skala rasio ini, komparasi atau

perbandingan dari tiap titik pada setiap unit pengukuran bersifat bebas. Angka-angka atau simbol numerik yang terdapat pada skala rasio dinyatakan dimulai dari angka nol, hanya saja unit-unit pengukurannya berbeda. Oleh karena itu, skala disebut “unique up to multiplication by a positive constant”, artinya perbandingan dua nilai dalam skala tidak akan mengalami perubahan bila seluruh angka dalam skala dikalikan dengan bilangan positif sehingga tidak akan mengubah maksud atau keterangan yang terkandung dalam skala tersebut. Contoh data rasio: tingkat kelulusan siswa SMA tahun 2020, tingkat kemiskinan di propinsi Sumatera Utara tahun 2021 dan lain sebagainya.

Jenis data kedua adalah data kualitatif. Data ini tidak dapat diukur dan di tuangkan dalam skala numerik/angka angka. Namun demikian, karena tradisi dalam statistika adalah bahwa semua data harus dalam bentuk numerik/angka maka. data kualitatif umumnya dikuantitatifkan agar formula statistika dapat dipergunakan. Caranya adalah dengan mengklasifikasinya dalam bentuk kategori. Pada dasarnya jenis data kualitatif dapat digolongkan menjadi:

- a. Data nominal, terjadi bila angka-angka atau simbol-simbol dipakai untuk menggolong-golongkan suatu objek, orang, atau suatu karakteristik. Dalam data nominal, penyusunan skala dalam-kelas merupakan suatu gugus atau rangkaian yang terpisah-pisah atau bebas. Satu-satunya hubungan yang terdapat di antaranya ialah sifat kesamaan, tiap anggota subkelas harus sama keadaannya untuk membentuk satu skala. Karena klasifikasi dalam skala nominal dinyatakan dengan jelas oleh suatu set daripada simbol-simbol maka skala nominal ini dikatakan “unique up to a one-to-one transformation”. Beberapa simbol yang menyatakan berbagai sub kelas dalam skala sesungguhnya dapat dipertukarkan, asalkan masih dapat dipahami. Simbol tentang jenis kelamin, tentang warna, musim misalnya dapat pertukaran. Sebagai contoh adalah pria di simbolkan sebagai angka 1, wanita sebagai angka 2 , musim kemarau angka 1 dan musim hujan di beri angka 2. Angka – angka simbolik tersebut bukan berarti nilai atau peringkat, melainkan pencirian atau pembeda

semata. Dalam hal ini berarti bahwa angka 1 berbeda dengan angka 2, dimana kemarau berbeda dengan penghujan.

- b. Data ordinal diberi batasan sebagai objek-objek dalam suatu kategori yang mungkin tidak berbeda dengan objek yang lain tetapi masing-masing objek tersebut tergabung dalam suatu hubungan dinyatakan dalam skala peringkat (Tabachnick & Fidell, 1996). Selanjutnya dapat dikatakan betapa suatu susunan skala ordinal akan memiliki karakter “unique up to a monotonic transformation” . Dalam diskursus ini, suatu hubungan yang mempunyai sifat yang selalu sama. Dalam data ordinal penskalaan dilakukan atas dasar kriteria yang terputus. Ilustrasinya adalah kelompok yang setuju dan tidak setuju, meskipun pembagian tersebut tidaklah tepat, karena konsep setuju tersebut masih bisa di bagi lagi menjadi lebih detail seperti sangat setuju dan setuju. Begitupun dengan tidak setuju, masih ada tingkatan sangat tidak setuju. Jadi ada semacam pemeringkatan atau tingkatan/derajat kualitas pada konsep tersebut.

5.2 Data Menurut Dimensi Waktu

Penggolongan data dalam perspektif ini di dasarkan pada dimensi waktu tertentu, yang dalam hal ini dapat dibedakan menjadi :

- a. Data *time series*, yaitu data yang disusun secara kronologis, artinya di konstruksi menurut urutan waktu pada suatu variabel tertentu. Data *time series* ini fungsinya adalah mendeteksi kemungkinan adanya perubahan dalam rentang waktu tertentu. Variasi adalah terjadinya variabel yang datanya bisa berubah antar waktu. Ilustrasinya adalah data harian, data mingguan hingga bulanan, atas pengunjung toko buku, perpustakaan daerah dan sebagainya.
- b. Data cross section, yaitu data yang dihimpun pada suatu titik/penggal waktu tertentu. Data cross section ditujukan untuk mengamati respons dalam kurun yang sama sehingga variasi terjadinya adalah antar pengamatan. Contoh: jumlah angka pengangguran pada setiap propinsi di Indonesia.
- c. Data pooling, merupakan bentuk kombinasi antara data *time series* dan *cross sektion* yang dipadukan sedemikian rupa. Misalnya, peneliti ingin mencermati jumlah lulusan perguruan tinggi negeri (akademi,

sekolah tinggi dan universitas) di kota Surakarta selama 3 tahun terakhir. Jumlah data cross section misalnya ada 3 yang terdiri dari 3 buah PTN (Poltekkes Kemenkes Surakarta, Institut Seni Indonesia, Surakarta dan Universitas Sebelas Maret Surakarta) maka observasi yang dilakukan adalah 3 PTN kali 3 tahun periode pengamatan, sehingga terjadi 9 kali pengamatan.

5.3 Data Menurut Cara Memperolehnya

Berdasarkan prosedur atau cara menghimpunnya, data dapat dibagi ke dalam:

- a. Data primer**, yaitu data yang dihimpun dan diolah serta di analisis sendiri oleh peneliti/observer secara langsung dari objeknya. Contohnya adalah seorang peneliti ingin mengetahui motivasi menggunakan media online instagram, maka dilakukan dengan wawancara langsung kepada informan yang telah ditentukan. Teknik pengumpulan data primer , sebagaimana dalam suatu penelitian pada umumnya, banyak dipengaruhi oleh rumusan permasalahan penelitian yang telah dibuat dalam rancangan riset. Maka dalam diskursus ini hal elementer bagi seorang peneliti bukan sekadar mengetahui bagaimana pelaksanaan penghimpunan data tersebut, namun juga memahami hal yang mempengaruhi teknik pengumpulan data tersebut. Harapannya adalah peneliti dapat menggunakan teknik pengumpulan data yang tepat. Selanjutnya, data primer dapat dihimpun atau diperoleh melalui:

- i. Observasi**

Teknik pengumpulan data dengan observasi (pengamatan) adalah suatu prosedur teknis atau katakana lah cara yang digunakan peneliti guna memperoleh data atau berbagai bentuk informasi yang berbentuk perilaku nonverbal informan atau responden, dengan tujuan agar dapat menjelaskan atau memberi jawaban atas pertanyaan penelitian sebagaimana yang telah dirumuskan sebelumnya. Pengamatan tingkah laku dilakukan sedemikian rupa sehingga informan atau responden tidak mengetahui jika sedang di amati, sehingga semua tingkah lakunya berjalan natural. Dalam teknik ini, dominasi peran dimiliki oleh peneliti, sebab informan hanyalah subjek – pasif saja, sehingga kualitas

riset akan sangat tergantung kepada tingkat ketelitian, kepekaan, kepekaan, dan pengendalian diri peneliti sendiri (Malo, 1986).

ii. Wawancara

Konsep wawancara adalah perbincangan yang sifatnya dialog/atau komunikasi dua arah, akan tetapi sumber inisiatifnya datang dari peneliti / pewawancara dengan tujuan mendapatkan berbagai informasi faktual dari informan/responden. Perbedaan peran dari pewawancara dan responden adalah jelas. Peneliti/pewawancara umumnya tidak saling mengenal sehingga harus memberikan situasi yang nyaman kondusif agar jawaban yang dilontarkan mencerminkan kondisi nyata atau faktual dan aktual berkenaan dengan tema penelitian. Jika dilakukan dengan sembarangan maka bukan tidak mungkin jawaban yang diberikan begitu datar, normatif atau apapun istilahnya, yang tidak mencerminkan kondisi yang sebenarnya. Jadi memang ada beberapa faktor yang memengaruhi dalam proses transfer informasi tersebut, yakni: peneliti /pewawancara itu sendiri, responden /informan, topik penelitian yang tertuang dalam daftar pertanyaan, dan situasi dan kondisi saat wawancara berlangsung.

iii. Kuesioner

Instrumen pengumpulan data lainnya adalah kuesioner. Instrumen ini merupakan alat penghimpun data data primer yang relatif efisien jika dibandingkan dengan teknik observasi dan *interview* atau wawancara. Dalam praktik, penelitian tidak lepas dari bea operasional pada berbagai tahapannya. Untuk tahapan pengumpulan data, kuesioner cenderung lebih rendah serapan biayanya. Pengumpulan data primer melalui kuesioner cukup mudah sebab tenaga lapangan 1 (satu) orang saja dapat mengerjakannya, ya tentu saja tergantung juga banyaknya responden dan tenggat waktu yang ditoleransi oleh tahapan penelitian itu sendiri.

Dalam kuesioner, peneliti memiliki lebih banyak waktu untuk mengumpulkan bukti, berbicara dengan orang lain, atau mempertimbangkan panjangnya jawaban dibandingkan dengan teknik wawancara. Dalam kuesioner umumnya dipandang lebih bersifat tidak mengacu pada orang tertentu, memberikan peluang kepada responden

untuk merahasiakan jati dirinya ketimbang cara-cara komunikasi lainnya. Kuesioner dapat berupa pertanyaan tertutup atau terbuka, dapat diberikan kepada responden secara langsung atau dikirim melalui pos/internet.

b. Data sekunder, yaitu data penelitian yang dihimpun dari bentuk data-data yang sudah jadi dalam berbagai bentuk karena telah diolah oleh orang lain. Dalam konteks ini data sekunder bisa di peroleh dari misalnya publikasi ilmiah, buku-buku ataupun jurnal-jurnal yang telah diterbitkan. Sebagaimana disebutkan di atas, bahwa ada sekunder diperoleh dari pihak lain setelah mengalami proses pengolahan. Seorang peneliti dapat mengakses data sekunder ini melalui berbagai sumber yang sifatnya kemudian bagi peneliti adalah sumber data sekunder. Saat ini dimana publikasi, jurnal, buku, monograph dan lainnya sangat mudah di akses secara digital, maka mengumpulkan data sekunder sangat mudah. Banyak lembaga riset, institusi pemerintahan, perguruan tinggi dan lembaga lain memiliki laman atau website, dimana publikasi berbagai hal di pajang disana. Dalam proses riset data sekunder digunakan untuk tiga buah tujuan yang elementer (Cooper& Emory, 1995).

- i. Tujuan pertama adalah untuk memenuhi kebutuhan akan acuan atau rujukan spesifik menyangkut hal-hal tertentu. Contohnya adalah jumlah penduduk suatu provinsi pada suatu waktu tahun tertentu atau tingkat pertumbuhan ekonomi rasional. Dari data skunder seorang peneliti dapat juga menelusuri acuan untuk menguji penemuan yang dilakukan oleh peneliti lain.
- ii. Tujuan yang kedua , bagaimanapun data data sekunder adalah penting bagi konstruksi penelitian yang lebih besar. adalah sebagai sebuah bagian terpadu dari sebuah studi penelitian yang lebih besar. Prosedur penelitian ini secara khas memerlukan data awal untuk menentukan apakah penelitian sebelumnya juga memberikan kontribusi pada riset yang sedang berjalan. Dalam bahasa lain dapat di katakan bahwa penelitian terdahulu ini merupakan acuan untuk menemukan *novelty* atau kebaruan yang mungkin bisa digali dengan seksama.
- iii. Tujuan ketiga data sekunder adalah untuk dijadikan sumber tunggal bagi penelitian yang sedang dilakukan ini. Dalam hal ini, ilustrasi yag

dapat diberikan adalah penelitian dalam bidang sejarah. Sebagaimana diketahui bahwa dalam penelitian sejarah data yang akan menjadi bahan untuk dianalisis seluruhnya menggunakan data-data sekunder dari peristiwa yang terjadi pada masa lampau. Dalam kajian yang sifatnya restropektif selalu memanfaatkan kajian-kajian yang telah terdokumentasikan pada masa lalu. Penelitian ini sebenarnya tetap mempunyai bobot ilmiah yang memadai, sebab bahan dasarnya yang berupa data sekunder juga merupakan hasil kajian sebelumnya yang juga menggunakan kaidah ilmiah.

5.4 Penyajian Data

Sebuah realitas empiris tidak akan berguna jika hanya di dikumpulkan, diolah, dan dianalisis tanpa kemudian disajikan dengan sistematis kepada pihak yang berkepentingan, misalnya pemerintah, lembaga swasta hingga masyarakat luas. Oleh karena itu perlu ditampilkan dalam model dan bentuk yang mudah dibaca dan dipahami oleh para pembuat keputusan yang kelak akan digunakan sebagai pijakan pembuatan keputusan. Bentuk yang ditampilkan bisa dalam tabel, gambar, grafik dan diagram.

Selanjutnya, untuk memudahkan memahami sajian data hasil kajian, pada umumnya data akan tampilkan dalam dua buah model utama, yakni tabel atau gambar dan grafik atau diagram.

a. Penyajian data dalam Bentuk Tabel

Apabila peneliti ingin menampilkan data yang telah diolah dan dianalisisnya itu kepada *user*, stake holder atau publik sekalipun, maka dalam penyajiannya pada bentuk tabel perlu memperhatikan:

1. Tabel hendaknya ada identitasnya, yakni diberi judul sesuai isi tabel, dibuat sederhana, ringkas dan jelas.
2. Tabel, apabila jumlah baris atau kolom cukup banyak sebaiknya diberi nomor urut, satuan hasil pengukuran juga harus diperhatikan.
3. Andai dibutuhkan penulisan data ada baiknya dibuat dalam kelompok-kelompok. Jika ada tujuan untuk komparasi maka hal yang perlu diperhatikan adalah letak, maka kelompok data yang akan dibandingkan diposisikan berdekatan.
4. Hal yang tak kalah elementer adalah pencantuman sumber data yang akan memberikan informasi bagi pembaca, darimana data didapatkan.

Tabel menurut perspektif pengaturan datanya, dapat diidentifikasi menjadi beberapa bentuk yang berbeda, yakni tabel baris-kolom, tabel kontingensi, dan tabel distribusi frekuensi.

i. Tabel baris-kolom

Tabel baris kolom adalah tabel yang menunjukkan atau memuat pengelompokan data sesuai perinciannya dalam bentuk baris dan kolom. Tabel baris-kolom dapat berupa tabel-baris kolom tunggal dan ganda.

Contoh :

Tabel 1. Kriteria UMKM dan Usaha Besar Berdasarkan Aset dan Omzet

Ukuran Usaha	Kriteria	
	Aset (tidak termasuk tanah & bangunan tempat usaha)	Omzet (dalam 1 tahun)
Usaha Mikro	Maksimal Rp 50 juta	Maksimal Rp 300 juta
Usaha Kecil	Lebih dari Rp 50 juta - Rp 500 juta	Lebih dari Rp 300 juta - Rp 2,5 miliar
Usaha Menengah	Lebih dari Rp 500 juta - Rp 10 miliar	Lebih dari Rp 2,5 miliar - Rp 50 miliar
Usaha Besar	Lebih dari Rp 10 miliar	Lebih dari Rp 50 miliar

Sumber: UU No.20/2008 tentang Usaha Mikro, Kecil, dan Menengah

ii. Tabel Kontingensi

Tabel ini menampilkan konstruksi data yang sesuai dengan rinciannya. Artinya, jika pada bagian baris tabel terdapat **m** baris dan bagian kolom terdapat berisikan **n** kolom maka didapatkan tabel kontingensi berukuran **m×n**.

Contoh :

Tabel 2. Nilai Ekspor-Impor Nasional Tahun 2020

No.	Bulan	Nilai Ekspor (US \$)	Berat Ekspor (KG)	Nilai Impor (US \$)	Berat Impor (KG)
1	Januari	13 636 412 653,62	50 900 144 920,20	14 268 720 284,00	12 143 552 871,00
2	Februari	14 042 089 243,23	49 671 209 459,93	11 548 100 132,00	13 059 584 031,00
3	Maret	14 031 292 077,89	54 068 776 030,33	13 352 176 374,00	14 432 183 255,00
4	April	12 159 824 545,01	45 172 672 643,65	12 535 233 221,00	15 051 046 827,00
5	Mei	10 452 625 424,79	42 236 217 598,43	8 438 627 383,00	10 084 426 066,00
6	Juni	12 006 813 612,70	46 387 784 702,64	10 760 317 981,00	11 505 323 617,00
7	Juli	13 689 902 558,81	46 069 923 524,57	10 464 299 676,00	11 323 464 029,00
8	Agustus	13 055 281 120,71	43 565 183 268,24	10 742 407 847,00	11 839 690 849,00

9	September	13 956 176 571,88	43 933 093 550,32	11 570 104 770,00	12 883 228 618,00
10	Oktober	14 363 443 251,43	47 117 845 245,64	10 786 016 684,00	12 091 546 513,00
11	November	15 258 422 043,10	53 282 636 840,90	12 664 414 194,00	12 137 149 554,00
12	Desember	16 539 555 059,63	57 272 738 271,50	14 438 376 084,00	15 330 669 376,00

Sumber : Badan Pusat Statistik, 2021

iii. Tabel Distribusi Frekuensi

Untuk tabel jenis ini, didefinisikan sebagai tabel yang menampilkan jumlah kejadian atau frekuensi yang berhasil secara empiris diamati.

Contoh :

Tabel 3 Hasil Ulangan Mahasiswa Fakultas X

Nilai	Jumlah Siswa
50-54	6
55-59	8
60-64	14
65-69	18
70-74	24
75-79	14
80-84	9
85-89	5
90-94	2
Jumlah	100

Sumber : Data Olahan Penulis

b. Penyajian dalam Bentuk Diagram

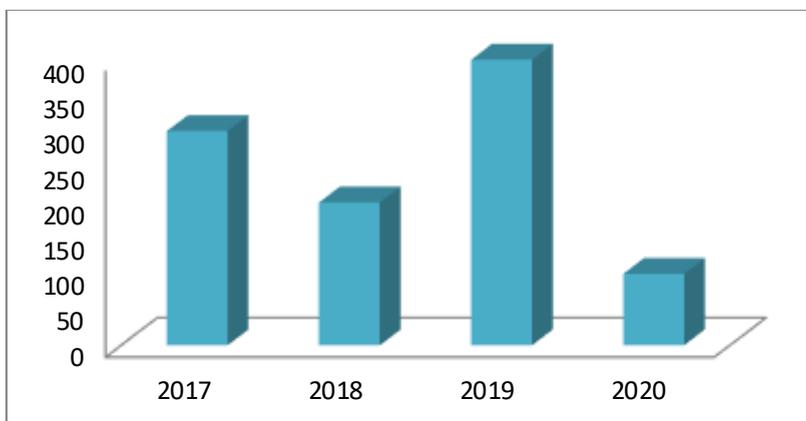
Bentuk penyajian data yang lain adalah diagram, yang tidak lain adalah bentuknya adalah visual atau gambar-gambar. Dalam bentuk penyajian ini ada variasi yang dapat dibuat yakni diagram batang, diagram pastel atau diagram lingkaran, diagram peta atau kartogram, diagram pencar atau diagram titik.

i. Diagram Batang

Bentuk diagram batang merupakan grafik data berupa persegi panjang yang diberi skala atau ukuran tertentu sesuai dengan sumber data yang telah digali. Dalam penggambarannya, batang tidak boleh saling berimpit dengan gambar batang yang lainnya. Penggambaran yang ideal adalah bahwa setiap batang berjarak yang sama.

Contoh :

Diagram 1: Panen Jagung Di Propinsi X dalam Ribu Ton



Sumber: Ilustrasi Penulis

ii. Diagram Lingkaran

Diagram lingkaran merupakan grafik data yang bentuknya adalah lingkaran. Lingkaran tersebut tidak utuh, karena telah dibelah menjadi bagian-bagian tertentu, sebagai representasi data yang telah dikumpulkan dan diolah tersebut. Pada umumnya belahan atau bagian diberi ukuran dalam persen.

Contoh:

Diagram 2 Laporan Penjualan



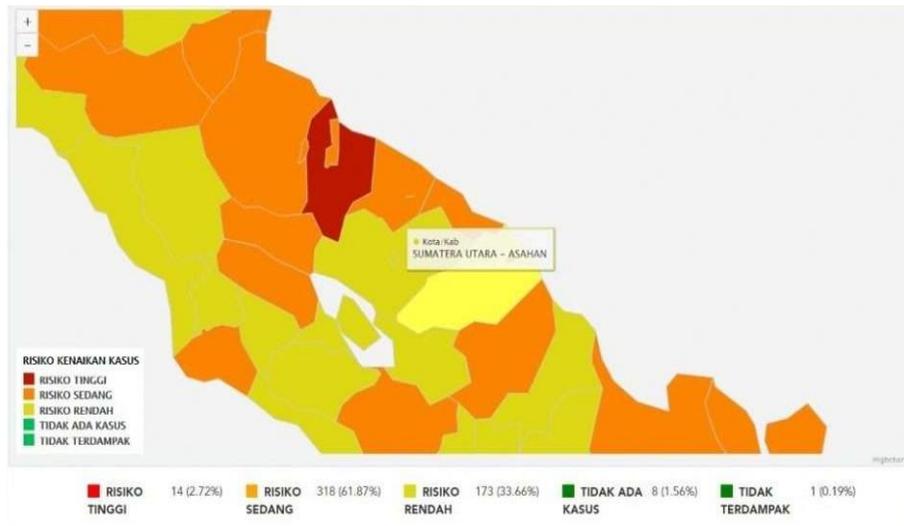
Sumber: Ilustrasi Penulis

iii. Diagram Peta atau Kartogram

Dalam bahasa visual, diagram yang berbentuk peta disebut dengan kartogram yang merepresentasikan data – data tertentu.

Contoh :

Diagram 3. Peta Kasus Covid-19 Pada Juli 2021



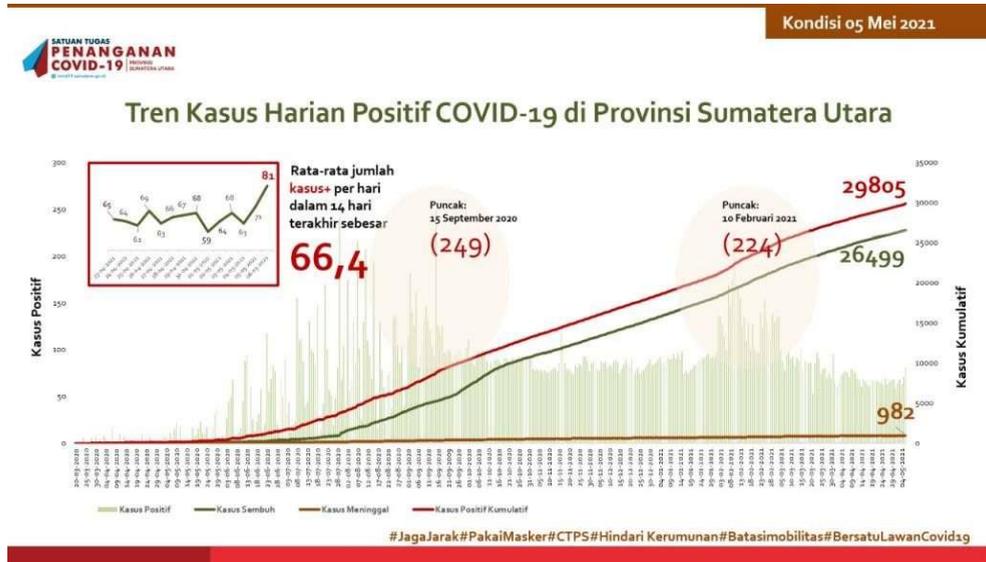
Gambar Kartogram : Sumber Pemda Asahan, 2021

iv. Diagram Garis

Grafik garis adalah grafik data berupa garis untuk menggambarkan perkembangan (tren) dari suatu data, misalnya data pandemi penyakit, perkembangan produksi, data perkembangan jumlah penduduk, data penjualan suatu barang, dan sebagainya.

Contoh:

Diagram 4. Visualisasi Garis Kasus Covid-19 Sumatera Utara



Gambar Kartogram : Sumber Pemda Asahan, 2021

SOAL SOAL LATIHAN

1. Ilmu statistika dan riset, terutama riset – riset kuantitatif tidak dapat dipisahkan. Deskripsikan bagaimana hubungan antara keduanya!
2. Penelitian tentu membutuhkan data – data yang akan dianalisis untuk memberikan jawaban atas *research question* yang ada. Tolong berikan penjelasan mengenai jenis data menurut cara memperolehnya, dan berikan contoh yang relevan!
3. Dalam statistika ada beberapa bentuk penyajian data dalam bentuk diagram. Sebutkan jenis diagram yang anda ketahui!
4. Apa yang saudara ketahui tentang statistika parametrik dan nonparametrik, jelaskan dengan seksama!
5. Pada dasarnya rumusan dalam statistika berlaku ketika data – data yang diperoleh adalah data kuantitatif. Pengamatan sebagai cara memperoleh data yang dilakukan ternyata mempertemukan saudara dengan realitas kualitatif. Jelaskan pendapat saudara!

BAB II

DISTRIBUSI FREKUENSI

2.1 Batasan Distribusi Frekuensi

Seorang peneliti belum mampu menyimpulkan sesuatu, apalagi kesimpulan mendalam atas data – data yang telah dihimpun, karena data masih dalam bentuk data mentah. Untuk memahaminya dibutuhkan cara – cara yang sistematis, yakni menyusun data dalam kelas – kelas tertentu, yang sering disebut sebagai distribusi frekuensi atau tabel frekuensi. Dengan kata lain batasan dari distribusi frekuensi adalah susunan data-data menurut kelas-kelas interval tertentu atau menurut kategori tertentu yang telah ditetapkan kedalam sebuah daftar atau tabel.

2.2 Membuat Tabel Distribusi Frekuensi

Untuk membuat tabel distribusi frekuensi diperlukan data mentah yang berasal dari data observasi mengenai hal tertentu yang berasal dari responden atau subjek penelitian. Data dapat berupa apa saja sepanjang menjadi bahan untuk menjawab pertanyaan penelitian. Dalam konteks ini dapat dicontohkan misalnya adalah data – data mengenai panen padi di suatu provinsi tertentu dalam kurun waktu lima tahun terakhir, jumlah penjualan produk suatu perusahaan perakitan mobil, jumlah produksi barang yang dihasilkan perusahaan tertentu dan lain sebagainya.

Contoh berikut memperlihatkan profil nilai ujian 80 orang mahasiswa fakultas X untuk mata kuliah statistika .

78	98	87	79	74	49	48	81
80	82	93	91	76	90	84	80
70	74	56	81	64	92	71	99
85	83	93	68	51	86	72	65
68	66	73	74	73	83	35	90
72	67	75	90	77	73	93	92
97	88	91	81	72	61	91	80

71	74	59	80	95	70	74	70
60	89	67	60	82	83	63	63
88	79	75	76	70	88	63	76

Berikut ini adalah ilustrasi untuk belajar membuat tabel distribusi frekuensi dengan panjang kelas interval yang sama. Adapaun beberapa langkah yang dapat dilakukan oleh peneliti adalah;

1. Tahap pertama adalah menentukan besaran atau nilai jarak, dimana dapat dilakukan dengan mengurangkan nilai data terbesar dengan yang terkecil. Dalam ilustrasi data mentah diatas dapat dilihat bahwa data terbesar adalah bernilai 99 dan data terkecil adalah 35, sehingga dapat dihitung nilai jarak yakni, $99 - 35 = 64$.
2. Tahapan kedua adalah menentukan jumlah kelas interval yang dibutuhkan peneliti untuk malakukan perhitungan statistik kelak. Dalam kajian statistika, jumlah kelas atau interval yang lazim diambil paling sedikit 5 tingkatan dan paling banyak 15 tingkatan kelas, dipilih menurut kebutuhan. Namun ada cara lain yakni menggunakan rumus Sturges, yaitu :

$$\text{Jumlah kelas} = 1 + 3,3 \log n$$

n menyatakan banyak data dan hasil akhir dijadikan bilangan bulat, untuk contoh di atas $n = 80$

Rumus Sturges di atas pada dasarnya memiliki kekurangan/kelemahan andai digunakan sebagai panduan/pedoman mutlak untuk menentukan jumlah kelas. Penggunaan rumus tersebut untuk n yang terlampau besar atau terlampau kecil dapat memberi hasil yang menyesatkan. Dalam praktik penggunaan rumus Sturges belum tentu menghasilkan jumlah kelas yang aplikatif untuk membantu penyusunan distribusi frekuensi. Bila berpedoman pada rumus Sturges guna menghitung jumlah kelas yang seharusnya digunakan dalam penyusunan distribusi maka hasilnya menjadi:

$$\text{Jumlah kelas} = 1 + 3,3 \log 80$$

$$\text{Jumlah kelas} = 1 + 3,3 (1,9031) = 7,28202$$

Tabel dapat dibuat dengan banyak kelas 7 atau 8 buah

- Tahapan berikutnya adalah menentukan panjang kelas interval (p). Dengan menggunakan asumsi atau cara perkiraan ditentukan oleh aturan berikut:

$$p = \frac{\text{Nilai jarak}}{\text{Banyak Kelas}}$$

Nilai p diambil sesuai dengan tingkat ketelitian atau daya presisi satuan data yang digunakan. Apabila data yang dihimpun berbentuk satuan, ambil nilai p teliti hingga level satuan. Begitu juga jika secara numerik yang untuk data hingga satu desimal, p ini juga diperhitungkan sampai satu desimal. Ilustrasi di atas, jika banyak kelas diambil 7 maka diperoleh:

$$p = \frac{64}{7}; \text{ dapat diambil } p = 9 \text{ atau } p = 10$$

- Selanjutnya adalah memilih ujung terbawah nilai interval pertama. Hal ini dapat ditempuh dengan mengambil sama dengan data terkecil atau nilai data yang lebih kecil dari data terkecil. Namun demikian, selisihnya harus kurang dari panjang kelas yang telah ditentukan. Selanjutnya tabel diselesaikan dengan menggunakan harga-harga yang telah dihitung.

Tabel. Nilai Ujian Statistika Deskriptif untuk 80 Orang Mahasiswa

Nilai ujian	Banyak Mahasiswa
31 – 40	1
41 – 50	2
51 – 60	5
61 - 70	15
71 – 80	25
81 – 90	20
91 – 100	12

2.3 Distribusi Frekuensi Relatif dan Kumulatif

Dalam tabel di atas, frekuensi dinyatakan dengan banyak data yang terdapat dalam tiap kelas; jadi dalam bentuk absolut. Jika frekuensi dinyatakan dalam persen; maka diperoleh daftar distribusi frekuensi relatif. Apabila frekuensi absolut dan relatif disajikan dalam sebuah tabel, akan tampak sebagai berikut:

Tabel Nilai Ujian Statistika Deskriptif untuk 80 Orang Mahasiswa

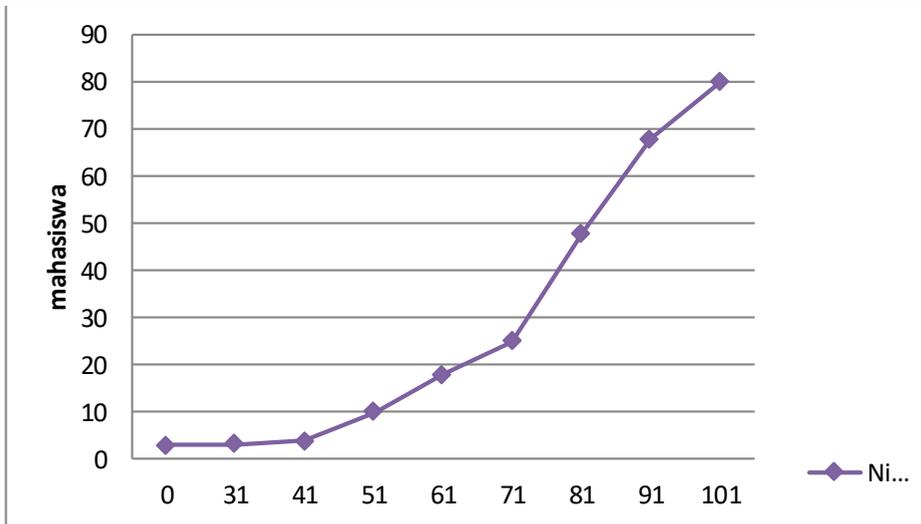
NilaiUjian	$f_{absolute}$	$f_{relatif}$
31 – 40	1	1,25
41 – 50	2	2,50
51 – 60	5	6,25
61 – 70	15	18,75
71 – 80	25	31,25
81 – 90	20	25,00
91 – 100	12	15,00
Jumlah	80	100,00

Ada lagi sebuah tabel yang dapat dibuat sehubungan dengan data diatas, yaitu tabel distribusi kumulatif. Tabel distribusi kumulatif dapat dibentuk dari tabel diatas dengan jalan menjumlahkan frekuensi demi frekuensi. Dikenal dua macam distribusi frekuensi kumulatif, yaitu “ kurang dari (less than) dan/atau lebih (more than)”. Tentu untuk kedua hal ini terdapat pula frekuensi-frekuensi absolut dan relatif. Untuk frekuensi kumulatif kurang dari (less than) dan atau lebih (more than) masing-masing dapat dilihat pada tabel dibawah ini.

Tabel Distribusi Kumulatif Kurang Dari

Distribusi Frekuensi absolut		Distribusi Frekuensi Kumulatif Kurang Dari	
Nilai nilai	Frekuensi	Nilai	Frekuensi Kumulatif
		Kurangdari 31	= 0
31 – 40	1	Kurangdari 41	0 + 1 = 1
41 – 50	2	Kurangdari 51	0 + 1 + 2 = 3
51 – 60	5	Kurangdari 61	0 + 1 + 2 + 5 = 8
61 – 70	15	Kurangdari 71	0 + 1 + 2 + 5 + 15 = 23
71 – 80	25	Kurangdari 81	0 + 1 + 2 + 5 + 15 + 25 = 48
81 – 90	20	Kurangdari 91	0 + 1 + 2 + 5 + 15 + 25 + 20 = 68
91 – 100	12	Kurangdari 101	0 + 1 + 2 + 5 + 15 + 25 + 20 + 12 = 80

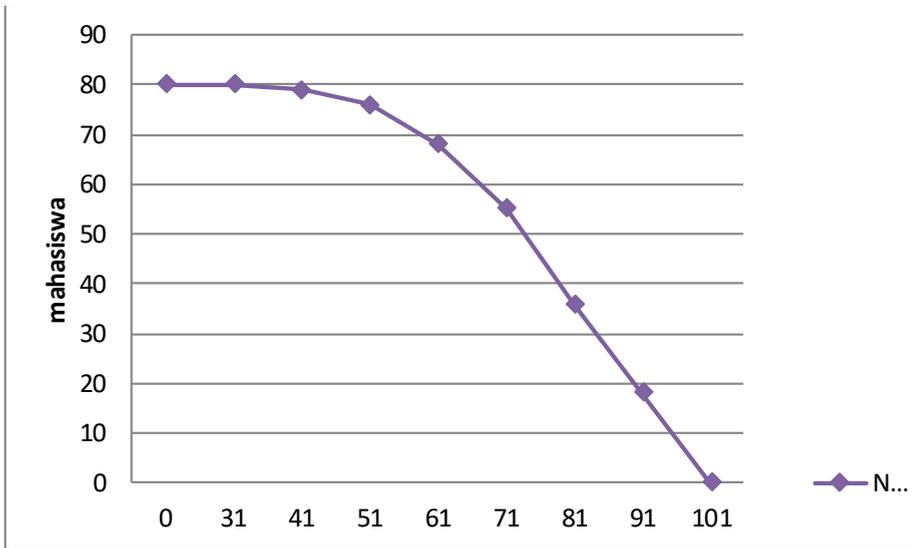
Grafik distribusi kumulatif kurang dari disebut ogive kurang dari atau ogive positif.



Tabel Distribusi Kumulatif atau Lebih

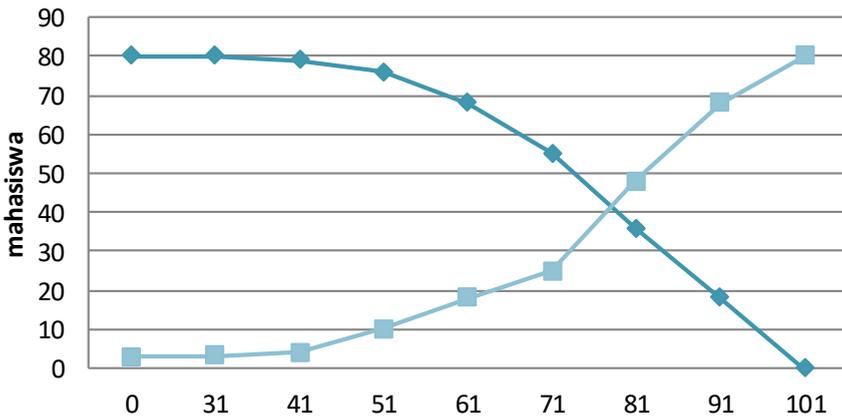
Distribusi Frekuensi absolut		Distribusi Frekuensi Kumulatif Kurang Dari	
Nilai-nilai	Frekuensi	Nilai	FrekuensiKumulatif
		31 atau lebih	= 80
31 – 40	1	41 atau lebih	80 – 1 = 79
41 – 50	2	51 atau lebih	80 – 1 – 2 = 77
51 – 60	5	61 atau lebih	80 – 1 – 2 – 5 = 72
61 – 70	15	71 atau lebih	80 – 1 – 2 – 5 – 15 = 57
71 – 80	25	81 atau lebih	80 – 1 – 2 – 5 – 15 – 25 = 32
81 – 90	20	91 atau lebih	80 – 1 – 2 – 5 – 15 – 25 – 20 = 12
91 – 100	12	101 atau lebih	80 – 1 – 2 – 5 – 15 – 25 – 20 – 12 = 0

Grafik distribusi kumulatif atau lebih disebut ogive atau lebih atau ogive negatif.



Gambar 2.2 Kurva Kumulatif atau lebih

Apabila kedua kurva ogive tersebut disatukan gambarnya akan terlihat sebagai berikut:



Gambar Kurva Distribusi Kumulatif Kurang Dari dan / atau lebih

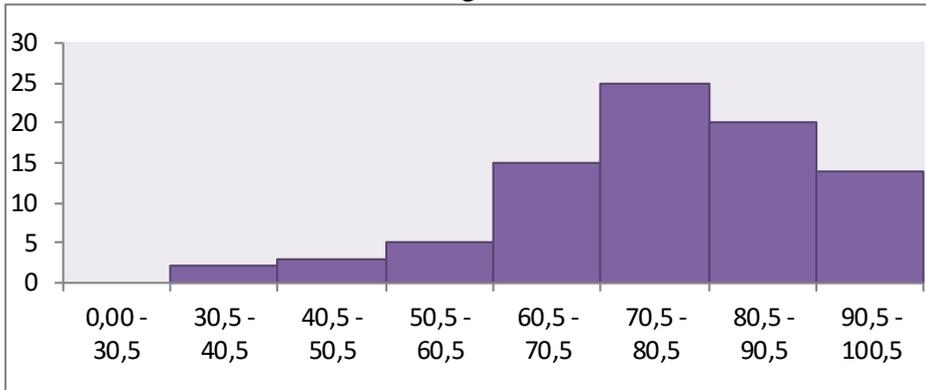
Kurva ogive diatas pada dasarnya memiliki beberapa kegunaan yang patut kita perhatikan (Dajan, 2000):

1. Kurva ogive umumnya digunakan untuk menggambarkan kumulasi frekuensi. Bila kita berkeinginan memperoleh keterangan mengenai jumlah mahasiswa yang memperoleh nilai diatas dan dibawah skor tertentu. Kurva ogive diatas lebih berguna jika dibandingkan dengan kurva frekuensi biasa.
2. Keragu-raguan dalam pemasukan angka-angka dalam kelas yang tertentu yang mungkin timbul karena persoalan kelas interval atau batas kelas tidak perlu ada.
3. Perhitungan statistika tentang median, kuartil, dan desil lebih mudah dilakukan dengan bantuan kurva ogive diatas.
4. Kurva ogive dapat digunakan bagi tujuan interpolasi. Misalnya, jumlah siswa yang mendapat nilai kurang dari 81 tetapi lebih dari 41 ternyata sebanyak $79 - 12 = 67$ orang. Hal sedemikian itu mudah sekali dilihat dari kurva ogive di atas.

2.4 Histogram Frekuensi

Penampilan data dalam bentuk diagram batang atau histogram diperlukan untuk mempermudah membaca realitas yang disampaikan lewat data tersebut. Pembuatan histogram berdasar data yang telah disusun tabel distribusi frekuensinya. Dalam praktik dapat digambarkan dengan membuat sumbu (X) mendatar untuk menyatakan besaran atau nilai frekuensi baik absolut maupun relatif. Angka yang dituliskan pada sumbu mendatar adalah batas-batas kelas interval dari nilai ujian. Bentuk diagram histogram frekuensi seperti diagram batang, hanya di sisi-sisi batangnya berimpitan.

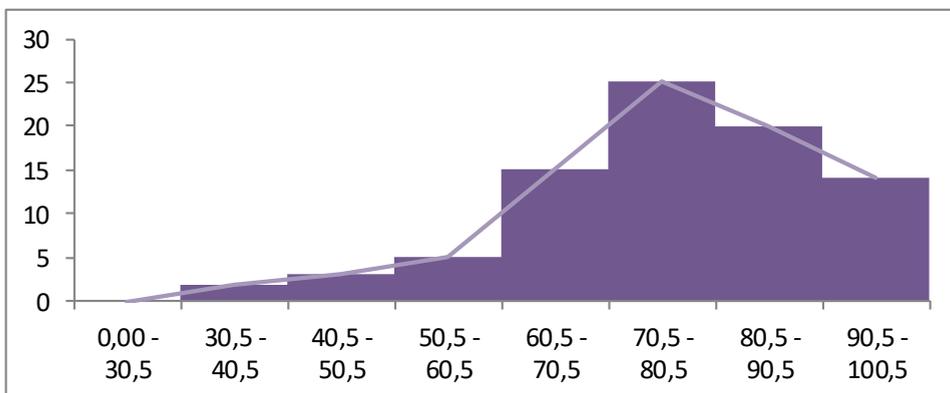
Gambar : Histogram frekuensi



Dalam bahasa lain, histogram sering kali dianggap sebagai grafik yang bertangga, mana salah satu fungsinya yang terpenting ialah menggambarkan perbedaan kelas-kelas dalam sebuah distribusi.

2.5 Poligon frekuensi

Dalam penggambaran yang lain, distribusi frekuensi dapat juga dinikmati dalam bentuk poligon frekuensi. Cara menggambarkan seperti ini sangat membantu peneliti atau para pembaca data andai ingin melakukan perbandingan antara dua atau beberapa distribusi frekuensi. Tahapan penggambaran poligon frekuensi pada umumnya dilakukan dengan jalan, pertama adalah menentukan titik tengah bagi tiap-tiap persegi panjang, setelah itu titik – titik itu dihubungkan dengan garis lurus utuh (atau bisa putus-putus). Untuk ilustrasi diatas, dpaat digambarkan sebagai berikut ini:



SOAL – SOAL LATIHAN

1. Deskripsikan dengan singkat hal yang harus diperhatikan dalam membuat distribusi frekuensi!
2. Deskripsikan jenis – jenis distribusi frekuensi yang saudara ketahui !
3. Deskripsikan apa yang dimaksud dengan nilai jarak, kelas interval, dan jumlah kelas !
4. Deskripsikan apa yang dimaksud dengan histogram frekuensi dan poligon distribusi, serta apa relasi antara keduanya?
5. Berikut ini adalah sebaran nilai mata kuliah akuntansi dasar 90 mahasiswa fakultas ekonomi universitas X.

70	89	78	70	84	87	79	82	83	91
75	79	75	53	51	46	71	59	67	79
66	51	47	84	52	70	90	87	59	48
75	61	60	76	82	83	40	34	95	53
83	81	67	44	85	64	52	98	87	93
90	80	61	88	73	84	86	81	78	46
67	82	85	66	91	65	82	90	93	90
66	54	85	76	79	78	73	60	51	32
61	58	89	78	87	69	82	85	56	55

Pertanyaan :

- a. Bagaimana tabel distribusi frekuensinya ? Gambarkan.
- b. Buatlah histogram frekuensi dan poligon distribusinya.

BAB III

KONSTRUKSI VARIABEL PENELITIAN DAN TEKNIK PENGUKURAN

3.1 Memahami Variabel Penelitian

Dalam seluk beluk penelitian kuantitatif, tahapan elementer yang harus dilakukan adalah menentukan variabel penelitian yang diperlukan. Dalam penelitian ini variabel-variabel penelitian harus ditentukan secara baik, sebagai titik perhatian riset. Lebih lanjut, variabel penelitian harus diterjemahkan dalam definisi operasional yang jelas dan harus dimanifestasikan dalam konsep – konsep yang bisa diukur. Kejelasan konsep itu dibutuhkan untuk kepentingan peneliti sendiri, agar mudah mengoleksi data – data empiris di lapangan.

Secara definisi ubahan atau variabel penelitian merupakan sebuah konsep yang memiliki variasi nilai. Variasi nilai ini jumlahnya tertentu atau minimal dua atribut, sehingga dapat untuk memberikan perbedaan. Variasi nilai dalam sebuah variabel penelitian ditetapkan oleh peneliti sebelumnya. Dalam bahasa lain dapat juga dikatakan bahwa variabel adalah konsep yang dapat diukur dan mempunyai variasi nilai.

Hal lain yang tak terpisahkan dari variabel adalah konsep. Penggunaan kata konsep dalam penelitian sangat penting untuk memberikan kejelasan dalam melakukan usaha pengukuran. Jadi memang konsep apa saja sepanjang mempunyai variasi nilai, maka merupakan variabel. Nah, sekarang konsep sendiri merupakan definisi yang digunakan dalam penelitian untuk mendeskripsikan atau mengabstraksikan fenomena sosial, ekonomi, politik, komunikasi atau yang lainnya yang sedang dikaji oleh peneliti.

Dalam penelitian, konsep yang mempunyai nilai yang bervariasi sudah pasti variabel, namun ada juga konsep yang sebatas konsep yang non variasi. Ilustrasinya adalah konsep mahasiswa , dimana konsep itu tidak memiliki variasi nilai sama sekali. Berbeda dengan konsep jenis kelamin mahasiswa, tinggi badan mahasiswa atau status perkawinan mahasiswa dimana konsep-konsep tersebut merupakan konsep yang memiliki variasi nilai. Konsep pendidikan juga bukan merupakan variabel, namun jenjang pendidikan dan jenis pendidikan adalah variabel.

Jika difokuskan lagi didalam variabel ada yang bentuknya adalah kategorial dan numerik. Jika mahasiswa adalah konsep yang tak mempunyai nilai, maka umur mahasiswa merupakan variabel yang mempunyai nilai numerik, yakni 18 tahun, 20 tahun 23 tahun dan sebagainya. Jenis kelamin mahasiswa juga merupakan variabel dalam penelitian, yakni masuk sebagai variabel kategorial dimana ada kategori mahasiswa pria dan mahasiswa perempuan atau mahasiswi. Konsep pegawai juga bukan variabel, namun tingkat pendidikan pegawai adalah variabel yakni variabel ordinal. Contoh lain lagi adalah konsep belajar, konsep itu bisa diadaptasi sebagai variabel dengan mengubahnya menjadi prestasi belajar, cara belajar dan sebagainya.

3.2 Karakteristik Variabel Penelitian

Dalam dunia riset (kuantitatif) variabel sesungguhnya menampilkan tiga ciri atau karakter yang fundamental. Pertama adalah bahwa variabel selalu atau dipersyaratkan mempunyai variasi nilai. Kemudian yang kedua adalah membedakan sebuah objek dengan objek yang lain dalam satu populasi. Ciri membedakan inilah yang sanggup mengidentifikasi antar objek riset yang dikaji. Ciri ketiga adalah *measurable* atau dapat di beri notasi ukuran.

Hal penting yang perlu dipahami dalam ciri variabel penelitian ini adalah ketika variabel merupakan konsep yang mempunyai variasi nilai dalam populasi atau sampel penelitian. Dalam hal ini dapat kita contohkan adalah indeks prestasi mahasiswa yang pastinya mempunyai nilai yang berbeda antara mahasiswa satu dengan yang lainnya. Jika kebetulan dalam populasi atau sampel itu indeks prestasi antar mahasiswa, maka tiada fungsi pembeda yang dipenuhi oleh variabel.

Ciri selanjutnya sebuah variabel membedakan antar objek dalam satu sampel atau populasi. Dalam hal ini dapat diidentifikasi misalnya adalah populasi mahasiswa yang dalam hal ini dianggap sama yakni populasi mahasiswa, dimana ada kesamaan status sebagai mahasiswa. Namun dalam ikatan kesamaan populasi itu, terdapat hal – hal yang saling berbeda satu sama lain misalnya hobi, makanan kegemaran, tinggi badan , umur, asal daerah, agama dan lain sebagainya.

Ciri yang ketiga bahwa variabel harus dapat diukur dalam notasi tertentu. Dalam hal ini dapat dicontohkan misalnya adalah mengukur variabel

hasil belajar, yakni bahwa hasil belajar dapat dinyatakan dalam angka – angka atau sering di sebut dengan nilai. Pada mata kuliah statistik, seorang mahasiswa A mendapat nilai 75, mahasiswa B mendapat nilai 90 dan mahasiswa C mendapat nilai 87 dan seterusnya.

Dalam perspektif lain, terdapat dua jenis variabel berdasarkan bulat atau tidaknya notasi angka atau nilai numerik yang diberikan kepada variabel. Dalam hal ini variabel dapat dibedakan menjadi dua yakni variabel diskrit dan kontinu. Secara definitif, variabel diskrit adalah variabel yang besarnya tidak mempunyai nilai pecahan, dalam konteks ini ilustrasinya adalah seperti jumlah pria yang berprofesi sebagai akuntan publik, jumlah armada taksi seorang pengusaha jasa transportasi, dan sebagainya. Selanjutnya, adalah variabel kontinu yakni merupakan variabel yang memiliki nilai pecahan, seperti misalnya adalah gaji karyawan, hasil tangkapan ikan nelayan, hasil panen apel petani di kabupaten Malang, dan sebagainya.

3.3 Relasi Antar Variabel

Pada umumnya dalam penelitian, realitas yang diteliti akan menampilkan satu hingga beberapa variabel. Maka itu penelitian dapat diselenggarakan dengan mengamati atau mengukur variabel yang jumlahnya lebih dari satu (univariat), mengamati dan mengukur dua variabel (disebut bivariat), atau mengamati dan mengukur lebih dari dua variabel (multivariate). Dalam klasifikasi jenis riset model penelitian-penelitian yang melibatkan dua variabel atau lebih, maka masuk dalam penelitian yang relasional atau mencari hubungan (sebab-akibat) antara variabel – variabel tersebut. Bisa juga mencari pengaruh dan menguji hipotesis – hipotesis yang telah dirumuskan sebelumnya. Dalam konteks hubungan antar variabel, maka variabel dalam penelitian dapat dibedakan menjadi beberapa jenis sebagai berikut :

a. Variabel Bebas (Independent Variable).

Variabel bebas, adalah variabel yang memengaruhi atau menjadi penyebab/kausa terjadinya perubahan pada variabel yang lain dalam konteks riset tersebut. Sebagai contoh, penelitian dengan judul “Pengaruh motivasi belajar terhadap prestasi mahasiswa fakultas ekonomi universitas X di kota Medan”, motivasi belajar mahasiswa adalah variabel bebas. Prestasi belajar mahasiswa akan tergantung/dipengaruhi oleh motivasi

belajar para mahasiswa tersebut. Dalam konteks pemodelan persamaan struktural atau sering disebut sebagai SEM, variabel motivasi belajar sebagai variabel bebas dibahasakan sebagai variabel eksogen.

b. Variabel Terikat (Dependent Variable)

Variabel ini merupakan variabel yang dipengaruhi oleh variabel bebas. Dengan kata lain variabel ini juga menjadi akibat karena adanya variabel bebas tersebut. Sebagai contohnya adalah bahwa “pengaruh motivasi belajar terhadap prestasi mahasiswa fakultas ekonomi universitas X di kota Medan” memiliki variabel terikat yakni prestasi mahasiswa ekonomi universitas X tersebut. Dalam kaitan ini prestasi mahasiswa dipengaruhi oleh motivasi belajar. Dalam perhitungan pemodelan SEM (Structural Equation Modeling), variabel yang tergantung atau dipengaruhi oleh variabel bebas disebut dengan variabel endogen.

c. Variabel Kontrol (Control variable)

Variabel jenis ini, merupakan salah satu bentuk variabel yang dalam hal tertentu atau dikendalikan pengaruhnya sehingga tidak mempunyai efek terhadap gejala – gejala yang diteliti. Artinya pengaruh variabel bebas terhadap variabel terikat tidak dipengaruhi oleh kondisi penentu luar atau faktor luar yang tidak masuk dalam wilayah penelitian.

Untuk lebih memahami mengenai variabel kontrol ini, berikut ini adalah gambaran mengenai relasi beberapa variabel yang akan mencerminkan fungsi dan arti variabel kontrol tersebut.

Contoh: pada penelitian berjudul: “Pengaruh metode mengajar terhadap prestasi mahasiswa pada mata kuliah statistika fakultas ekonomi universitas X” Pada desain penelitian itu, variabel bebasnya adalah metode mengajar, variabel terikatnya prestasi belajar mahasiswa, dan variabel kontrolnya adalah mata kuliah yang sama, yang dalam konteks ini adalah mata kuliah statistika. Dapat dipahami selanjutnya adalah variabel kontrol itu akan memberikan informasi bahwa besarnya pengaruh metode mengajar terhadap prestasi mahasiswa dapat diketahui dengan lebih pasti.

d. Variabel Moderator (Moderator Variable).

Variabel dalam kategori ini, pada penelitian atau pada hubungan antar variabel akan memperkuat atau sebaliknya, akan memperlemah, pengaruh variabel bebas terhadap variabel terikat.

3.4 Pengukuran Variabel Penelitian

Konstruksi pengukuran terhadap variabel penelitian, pada dasarnya memuat empat bentuk yang masing–masing mempunyai peruntukan yang berbeda – beda, sesuai dengan data yang akan diukur. Pada umumnya dapat dibedakan menjadi empat strata sebagai berikut ini:

a. Skala Nominal

Skala nominal digunakan untuk mengukur data – data nominal. Data ini sesungguhnya hanya dapat dideferensiasikan semata dan tidak untuk diurutkan dan juga dibandingkan dengan yang lain, sehingga pemakaian angka hanya sebatas notasi simbolik saja. Sekadar contoh adalah data jenis kelamin (yang dibedakan sebagai pria dan wanita) , agama seseorang (Islam, Kristen, Katolik, Hindu dan Budha), nama fakultas (Ekonomi, FKIP, FISIP, Ilmu Budaya , dsb) jenis sekolah dan sebagainya. Dalam diskursus ini terlihat bahwa pada skala nominal variasi nilai dari variabel yang diukur tersebut merupakan kategori-kategori yang setara. Penggunaan skala nominal ini, memperlihatkan betapa variasi nilai dari variabel yang diukur tidak menampilkan adanya interval dan derajat/strata atau ranking. Pemakaian skala nominal ini menunjukkan bahwa variasi nilai dari variabel yang diukur tidak menunjukkan jarak (interval) ataupun tingkatan (rangking) antara kategori-kategori dalam pengukuran variabel tersebut.

b. Skala Ordinal

Penskalaan dalam hal ini mempertimbangkan adanya tingkatan atas data yang diamati. Ada aktivitas penjenjangan yang didasarkan ada data yang diamati, yang kemudian responden dapat diurutkan sesuai dengan karakteristik yang melekat pada responden itu sendiri. Pada umumnya data yang diukur akan memperlihatkan tingkatan yang berbeda, yakni mulai tingkatan paling rendah hingga tingkatan paling tinggi. Dapat dicontohkan seluruh civitas akademika kampus, maka ada pengelompokan misalnya mahasiswa, dosen, kaprodi, wakil dekan, dekan, wakil rektor, dan rektor. Pengelompokan dalam skala ordinal ini juga secara langsung menampilkan tingkatan, misalnya variabel sikap pasien terhadap layanan klinik / rumah sakit dibedakan menjadi tidak puas, puas dan puas sekali, variabel tingkat sosial ekonomi masyarakat dibedakan menjadi menjadi rendah, sedang, dan tinggi dan sebagainya. Pengukuran skala secara ordinal ini menampilkan

lambang dan urutan (tingkatan) tanpa adanya indikasi mengenai besaran atau jumlah absolut atribut tiap-tiap tingkatan tersebut, selain itu tidak menampilkan interval antar tingkatan.

c. Skala Interval

Dalam pengukuran, salah satu realitas data empirik dapat diukur dengan penskalaan interval. Data yang kemudian diukur dengan penskalaan interval adalah data yang mempunyai perbedaan, urutan dan juga jarak perbedaan dalam komposisi urutan tersebut, namun data tersebut tidak memiliki angka nol yang absolut. Dengan kata lain skala interval memberikan informasi mengenai kesenjangan perbedaan atau interval antara objek yang satu dengan yang lainnya. Jarak dalam skala interval unit – unit pengukuran dan titik nol tidak bersifat absolut, sehingga tidak ada data informasi yang mutlak dari atribut suatu objek pengukuran. Contohnya adalah nilai ulangan akhir semester mahasiswa, 3, 4, 5,6,7 dan 8. Jika dilihat, maka urutan angka antara 1 sampai dengan 10 memiliki berjarak satu (1) satuan pada setiap tingkatannya. Jarak satuan antara 4 dan 5, sama dengan 6 dan 7, sama dengan 8 dan 9. Meski demikian angka tersebut tidak memiliki perbandingan, semisal nilai 4 mahasiswa bukan berarti daya kemampuan akademiknya mutlak separohnya dari mahasiswa lain yang mendapat nilai 8. Hal itu karena angka angka itu sebagai data interval tidak bisa dibandingkan.

Contoh lain yang dapat di simak adalah misalnya akan mengukur temperatur larutan kimia dalam laboratorium, dengan skala suhu yang dibuat oleh Reamur dan Celcius. Unit pengukuran dan titik nol dari kedua skala itu sama (yang dikatakan sebagai titik beku air) namun titik akhir keduanya berbeda yakni Reamur 80 dan Celcius 100 derajat. (angka tersebut dianggap sebagai titik didih air dalam percobaan mereka).

Reamur	0	40	64	80
Celcius	0	50	80	100

Dengan demikian, kedua skala ini menampilkan informasi yang sama banyaknya dan sama jenisnya, karena kedua skala tersebut berhubungan secara linear, dan dapat saling ditransformasikan. Notasi angka pada skala Reamur terhadap Celcius adalah $R = \frac{4}{5} C$ dan sebaliknya Celcius terhadap Reamur adalah $C = \frac{5}{4} R$. Disini, rasio perbedaan suhu air tidak bergantung

pada unit pengukuran dan titik nol maupun titik didihnya. Misalnya, "membeku" terjadi pada 0 derajat skala Celcius dan "mendidih" terjadi pada 100 derajat Celcius. Sementara pada skala Reamur, "membeku" terjadi pada 0 derajat dan "mendidih" pada 80 derajat. Variasi suhu yang sama pada kedua skala tersebut ditunjukkan pada tabel diatas.

d. Skala Rasio

Skala rasio merupakan pengukuran atas data yang mempunyai perbedaan, urutan dan interval yang sama dalam konfigurasi data itu, serta yang terpenting adalah mempunyai nilai nol yang bernilai mutlak. Nilai nol dalam angka nol pada skala ini benar – benar merepresentasikan bahwa gejala atau realitas yang diukur itu benar benar tidak ada. Skala rasio diperoleh dari variabel rasio yang dapat diperbandingkan, terdapat tingkatan, jarak interval, dan mencerminkan perbedaan. Dalam skala rasio ini, perbandingan antar titik - titik skala tidak akan tergantung kepada unit pengukurannya. Sebagai ilustrasi, jika kita mengukur panjang suatu objek riset dengan skala rasio, maka skala centimeter dan inchi akan memiliki titik nol yang mutlak, yang sejati. Perbandingan antar dua panjang benda yang diukur tidak akan terikat kepada unit pengukurannya. Dengan demikian jika ditetapkan panjang dalam centimeter maupun inchi, maka rasio panjang benda yang diukur dalam satuan centimeter adalah sama dengan rasio antara panjang dalam inchi.

3.5 Teknik Skala dan Pengumpulan Data

Di dalam menyusun instrumen penelitian, seorang peneliti harus memahami mengenai jenis – jenis skala pengukuran yang akan digunakan. Tujuannya adalah agar data yang akan diukur sesuai dengan instrumen pengukurannya, sehingga hasil pengukuran bisa dipercaya serta *reliable* terhadap permasalahan (Riduawan, 2005).

Terdapat beberapa skala pengukuran yang dapat digunakan oleh peneliti untuk memberikan alat/instrumen yang tepat dalam pengukuran data di lapangan. Bentuk – bentuk skala yang harus dipahami dalam penelitian antara lain: (a) Skala likert, (b) Skala Guttman, (c) Skala Semantik Deferensial, (d) Rating Scala dan (e) Skala Thurstone.

a. Skala Likert

Penskalaan ini digunakan untuk mengukur sikap, pendapat atau persepsi seseorang/kelompok orang yang didudukkan sebagai responden penelitian, mengenai kejadian atau gejala sosial yang akan diteliti. *Skala likert* didisain untuk memberikan derajat penilaian subjek penelitian setuju atau tidak setuju dengan pernyataan yang dituangkan dalam pertanyaan penelitian.

Pernyataan	1	2	3	4	5
1. Ditengah kontroversi penggunaan vaksin untuk membantu imunitas, bahkan orang yang 2 kali di vaksin pun, ketika terjangkiti virus covid 19 ada yang meninggal dunia. Vaksin relatif tidak di perlukan sebab imunitas juga bisa dibangun dari faktor lain.					
2. Visi bisnis PT Kultura Digital Media ini , kedepan berorientasi mendorong tingkat kepuasan konsumen.					
3. Konsep marketing mix akan memberi kontribusi maksimum pada penjualan, mengalahkan teknik marketing yang sepenuhnya dilakukan dengan teknik digital secara total					

NB: 1 2 3 4 5
 Sangat tidak setuju tidak setuju netral setuju sangat setuju

b. Skala Guttman

Pengukuran dengan skala Guttman sifatnya adalah akumulatif, yang artinya ketika seorang responden menyisakan atau tak menjawab pertanyaan yang berbobot lebih berat, maka kecenderungan yang terjadi adalah responden akan mengiyakan pertanyaan yang dinilai kurang berbobot lainnya. Pada titik ini skala Guttman bisa dipahami sebagai skala yang mengukur satu dimensi saja dari variabel yang multidimensi. Dengan demikian menyusun pertanyaan dengan skala ini akan lebih baik diurutkan secara herarkis untuk melihat pernyataan sikap seseorang. Kecenderungan umum adalah jika responden menyatakan “tidak” terhadap sederetan

pernyataan sikap tertentu sebelumnya, maka ia akan menyatakan “lebih dari tidak” terhadap pernyataan berikutnya. Dengan demikian skala Guttman adalah skala yang digunakan untuk jawaban yang sifatnya jelas-tegas-pasti serta mempunyai bobot konsistensi yang tinggi. Jawaban informan yang masuk kategori ini misalnya, ya – tidak, yakin-tidak yakin, benar-salah, positif-negatif dan sebagainya. Contoh penerapan skala Guttman dalam instrument penelitian adalah sebagai berikut :

No	Pernyataan & Jawaban
1	Apakah anda yakin bahwa pergantian rektor pasti menyelesaikan persoalan di kampus ini ? Yakin Tidak yakin
2	Setujukah anda jika rektor saat ini turun dari jabatannya ? Setuju Tidak setuju
3	Pernakah menurut anda selama kepemimpinan rektor saat ini, yang bersangkutan membuat kebijakan kontroversial ? Pernah Tidak Pernah
4	Kepemimpinan rektor saat ini berdampak positif bagi lingkungan kampus pada masa milenial seperti sekarang ini. Benar Salah
5	Rektor sebagai penguasa anggaran di kampus selalu mengetahui mal administrasi dalam lingkungan kampus. Iya Tidak

c. Skala Deferenstial Semantik

Penskalaan jenis ini sebenarnya adalah upaya untuk mengakomodasi sikap atau persepsi responden terhadap gejala yang diteliti, dengan memberikan ruang sikap tertentu diantara dua buah sikap yang secara bipolar

berhadapan secara diametral. Misalnya dapat kita identifikasi betapa diantara dua titik panas dan dingin, ada ruang yang memberikan informan/responden untuk menampilkan persepsinya. Ruang tersebut adalah ruang semantik. Contoh yang dapat diberikan adalah sebagai berikut :

Berilah tanda X pada skala dibawah ini yang paling cocok dengan persepsi anda atas realitas yang terjadi.

a) Kekuatan pengawasan orang terhadap anak mereka yang telah berstatus

	5	4	3	2	1	
Ketat	x					Longgar
	5	4	3	2	1	
Kerat Dilakukan	x					Tidak Pernah
	5	4	3	2	1	
Lemah	x					Kuat
	5	4	3	2	1	
Aktif	x					Pasif

mahasiswa menyangkut pergaulan bebas.

d. Rating Scala

Dalam skala ini yang penting dipahami adalah bahwa data yang masuk dalam penskalaan rating adalah data – data mentah yang berupa angka yang kemudian ditafsirkan dalam perspektif kuantitatif. Dalam pembuatannya, skala rating ini memberikan pilihan kepada responden untuk memilih salah satu jawaban kuantitatif yang telah disediakan. Kelemahannya memang, misalnya 5 (lima) orang memilih jawaban/skor 3 pada sebuah pertanyaan, maka sebenarnya persepsi 5 (lima) orang belum tentu sama maknanya, meskipun dapat dilihat bahwa mereka sama – sama memilih angka 5. Contoh instrumen pengukuran dengan skala rating adalah:

Peneliti hendak mengetahui konsep penciptaan keluarga bahagia, maka berilah tanda “x” pada jawaban yang sesuai dengan pendapat anda.

No	Pernyataan Tentang Keluarga Bahagia	Interval Jawaban				
		SB	B	CB	KB	STB
		5	4	3	2	1
1	Implementasi Nilai Keagamaan	x				
2	Tingkat Pendidikan Anak		x			
3	Pengelolaan finansial		x			
4	Kesempatan melakukan wisata, liburan dll				x	
5	Bentuk pergaulan dengan sesama, lingkungan		x			
6	Peraturan internal keluarga		x			
7	Pola kebiasaan dan pandangan hidup keluarga		x			
8	Distribusi tanggungjawab pada anggota keluarga			x		
9	Relasi antar anggota keluarga				x	
10	Mata pencaharian kepala/ibu rumah tangga				x	

e. Skala *Thurstone*

Skala ini memberikan pilihan jawaban kepada responden yang sesuai menurut pilihan mereka, dari beberapa pernyataan yang menampilkan berbagai pandangan yang saling bervariasi/berbeda. Secara umum setiap item pernyataan bernilai dalam kisaran 1 hingga 10, atau mempunyai asosiasi 1 hingga 10. Contohnya adalah sebagai berikut :

Tes Rekrutmen Staf Pengajar Fakultas Ekonomi Universitas “X“. Tolong pilihlah 5 buah jawaban dari pernyataan dibawah ini sesuai dengan perspektif saudara.	
No	Pernyataan
1	Pilihan profesi saya adalah akademisi/staf pengajar karena mengabdikan pada pengembangan dan diseminasi ilmu, sebuah profesi mulia, berwibawa dan agung.
2	Sebagai mahasiswa ekonomi yang kritis, identitas sebagai mahasiswa jurusan ekonomi sebaiknya dipertegas, misalnya mengenakan atribut tertentu.
3	Rasanya sangat bangga bila menguasai ilmu lalu diajarkan di depan kelas, dihadapan peserta seminar, daripada menguasai ilmu tapi tidak diajarkan.
4	Menjadi dosen, pada taraf tertentu makmur, namun dosen di perguruan tinggi kecil, sebagaimana diberitakan di berbagai media, gajinya kecil, minim tunjangan, beban kerja berat, belum lagi berhadapan dengan aneka tingkah polah mahasiswa. Inilah yang memperlihatkan menjadi dosen itu makan hati.

5	Bangga rasanya, puas rasanya, dan bersyukur rasanya bila bisa berbagi ilmu, menjelaskan dengan runtut di hadapan mahasiswa, lalu mereka takjub.
6	Kadang saya bergumam, akulah manusia paling berguna, seorang dosen yang menjadi pewaris ilmu pengetahuan, mengajarkan ilmu pada generasi penerus, aktor intelektual kemajuan bangsa.
7	Gaji dosen itu seyogyanya lebih besar dari pegawai lain. Karena jangan sampai guyonan "dosen = di dodos metu saksen (dosen = diubek-ubek saku bajunya, Cuma ada 1 sen) terjadi nyata
8	Dosen yang tidak aktif, mengajar seenaknya, sering indisipliner tidak adahaksa sekali untuk bangga dan membanggakan perannya, ketika ada mahasiswa didiknya memenangkan olimpiade ekonomi nasional.
9	Pembimbingan oleh dosen kepada mahasiswa sebaiknya intensif, dan dapat dipertanggungjawabkan, sebab ilmu itu juga akan diajarkan kembali, ketika mahasiswa anak bimbing itu juga berkeinginan mengajar
10	Saya mahasiswa fakultas ekonomi, dan saya mempunyai bisnis pisang goreng di pinggir jalan sebagai penopang hidup, karena tidak disokong orang tua lagi. Demi wibawa dan nama besar universitas dimana saya kuliah, saya akan sembunyikan identitas sebagai mahasiswa.

Selanjutnya, dalam desain riset yang bercorak kualitatif, atau penelitian kuantitatif yang melibatkan data yang sifatnya kualitatif pun, maka poses pengukuran variabel yang bersifat kualitatif tidak mungkin secara langsung dilakukan. Pengukuran hanya dapat dilakukan dengan melalui tahapan – tahapan tertentu yang sudah dipersiapkan sebagai berikut :

- a. Peneliti menentukan jenis / bentuk simbol tertentu yang akan digunakan untuk mengukur variabel atau karakteristik yang akan dikaji.
- b. Peneliti menentukan dimensi atau indikator apa saja yang tepat akan digunakan sebagai alat pengukur dalam penelitian
- c. Peneliti menentukan skala yang akan digunakan pada variabel yang diukur dan juga alat pengukurnya itu sendiri,
- d. Peneliti dengan cermat menentukan jumlah skala, yakni untuk variabel yang akan diukur dan juga alat pengukurnya itu sendiri.
- e. Peneliti menentukan kriteria masing-masing skala penilaian, menjabarkan dimensi/indikator ke dalam item-item pengamatan.
- f. Kompatibel alat ukur dapat ditentukan dengan baik, yakni ketika peneliti melakukan uji validitas dan reliabilitas.

Penelitian bisa sampai pada tahapan akhir, yang memperlihatkan kesimpulan tentu melalui pengolahan / pengujian atas data yang penelitian yang telah di kumpulkan. Data data penelitian , dalam hal ini adalah data

primer, dapat diakses atau dihimpun melalui penyebaran kuisioner dan juga dengan wawancara langsung (baik wawancara bebas, terstruktur dan wawancara mendalam), dan dikuatkan oleh pengamatan/observasi secara langsung *on the spot*. Data utama yakni data primer tadi dalam tradisi riset perlu diperkuat dengan sumbangan data sekunder yang diakses dari berbagai literatur yang relevan.

Kemudian, data yang berhasil dihimpun dari lapangan, harus memenuhi setidaknya dua prinsip mendasar yang bersifat elementer :

1. *Data harus memenuhi prinsip reliabilitas dengan syarat sebagai berikut ini:*

- a. Data harus bersifat objektif, artinya sesuai dengan realitas yang faktual/sebenarnya, sesuai dengan apa adanya di lapangan.
- b. Data harus bersifat representatif, dimana data sampel itu harus dapat mewakili seluruh populasi dimana sampel di ambil.
- c. Data harus bersifat cermat dan teliti serta se-rigid mungkin, artinya hanya mempunyai penyimpangan standar yang cukup kecil. Kaidahnya adalah bahwa perkiraan parameter populasi akan dikatakan baik bila mempunyai derajat penyimpangan standar yang kecil.

2. *Data harus memenuhi prinsip validitas, dengan syarat sebagai berikut:*

- a. Data harus bersifat kekininian atau tepat waktu (up to date), artinya data yang dihimpun tidak boleh kadaluarsa. Hal itu untuk menjaga agar faktor penyebab data menjadi bias atau semu semakin kecil, sebab data yang bias atau semu akan menyebabkan analisis yang keliru.
- b. Data harus harus kontekstual dan relevan, artinya data yang dihimpun harus ada hubungannya dengan permasalahan yang diangkat dalam penelitian dan akan diselesaikan. Dalam bahasa metaforik, data tidak boleh asing dengan masalah.

Proses atau tahapan penelitian setelah data data yang diperlukan terhimpun dengan baik, maka di perlukan upaya untuk mengelola dan menganalisis dengan seksama agar didapatkan gambaran yang presisi mengenai realitas yang diteliti. Beberapa perlakuan yang wajib dilakukan oleh peneliti terhadap data yang telah didapatkan dari lapangan adalah *editing, coding, dan tabulasi*.

1) *Editing*

Aktivitas editing pada dasarnya adalah kegiatan yang cukup menuntut sifat yang teliti dan cermat. Proses ini adalah aktivitas mencermati dan memeriksa validitas serta reliabilitas data-data yang telah masuk ke peneliti. Termasuk dalam aktivitas ini misalnya adalah memeriksa dengan teliti kelengkapan pengisian daftar pertanyaan atau kuisisioner, memeriksa kejelasan makna dari jawaban responden, derajat relevansi jawaban, tingkat konsistensi dan koherensi antar jawaban dan melihat kesamaan/konsistensi penggunaan satuan pengukuran data.

2) *Coding*

Proses yang kedua adalah *coding* yang merupakan aktivitas memilah – milah atau mengklasifikasikan data/jawaban yang telah dihimpun menurut masing-masing kategori, sebagaimana yang telah ditetapkan. Pada tiap kategori jawaban yang berbeda akan dilabeli kode kode yang berbeda pula. Hal yang perlu dicermati dalam kegiatan ini adalah setiap kategori yang sama harus diberi kode yang sama, teknis klasifikasi yang lain adalah diantara kategori yang satu dan lainnya dibatasi atau dipisahkan dengan tegas, sehingga tidak tumpang-tindih.

3) *Tabulasi*

Perlakuan terhadap adat yang masuk berikutnya adalah tabulasi yang merupakan aktivitas *summary* atau meringkas data yang masuk (data mentah) ke dalam bentuk yang lebih komunikatif. Pola atau format yang komunikatif ini biasanya dalam bentuk tabel-tabel yang telah dipersiapkan sebelumnya. Aktivitas tabulasi ini ada beberapa tahapan, yang *pertama* adalah mempersiapkan format tabel dengan jumlah kolom dan baris diperhitungkan selaras dengan kebutuhan yang ada. *Kedua*, adalah mencacah atau menghitung besarnya frekuensi untuk setiap kategori jawaban yang masuk, dan *ketiga* menyusun data-data tersebut secara distributif atau dalam bentuk tabel frekuensi dengan maksud agar data yang sudah tersusun bagus dan rapi mudah untuk dibaca, dianalisis serta dipahami.

SOAL – SOAL LATIHAN

1. Apa yang dimaksud dengan ‘‘konsep’’ dan ‘‘variabel’’ dalam penelitian, dan apa perbedaan diantara keduanya. Jelaskan dan beri contoh.
2. Apa yang dimaksud dengan karakteristik variabel, jelaskan !
3. Penskalaan dalam pengukuran data ada beberapa macam. Tolong jelaskan dan beri contoh !
4. Apa yang dimaksud dengan prinsip data yang elementer ? Uraikan dengan jelas !
5. Pasca koleksi data dengan berbagai instrument, bagaimana kelanjutan dari pengelolaan data itu?

BAB IV

POPULASI, SAMPEL, DAN TEKNIK PENGAMBILAN SAMPEL

4.1 Populasi Dalam Riset

Realitas empirik tentu berada dalam entitas tertentu, yang dalam bahasa penelitian, entitas itu adalah domain penelitian atau lebih populer disebut dengan populasi. Dari konstruksi itu dapat dipahami bahwa populasi adalah keseluruhan objek riset (sosok, peristiwa, dan lainnya) dengan kriteria dan karakteristik tertentu. Objek riset ini bisa bersifat nyata, konkret (*tangible*) maupun objek-objek yang tak kentara, yang abstrak (*untangible*).

Populasi dapat dikatakan “kumpulan” banyak sampel penelitian, sehingga didalam penelitian sangat diperlukan penentuan sampel tersebut sebagai cara untuk “memudahkan” dalam membaca fenomena atau realitas yang ada. Dalam penelitian, banyaknya objek dalam suatu populasi disebut dengan ukuran/jumlah populasi, dan untuk mempermudah dalam pelaksanaan penelitian pada umumnya di beri notasi atau dilambangkan dengan huruf N. Jumlah atau ukuran suatu populasi yang dapat dicacah/dihitung (bersifat *countable*) disebut dengan populasi terhingga/berhingga, tidak mempertimbangkan seberapaapun besarnya. Berbalikan dengan itu, jumlah/ukuran populasi yang sangat besar, sehingga kita kesulitan sekali untuk melakukan penghitungan (bersifat *uncountable*), populasi itu merupakan populasi tak hingga atau tak berhingga

Populasi yang sangat besar namun masih dapat dicacah atau dihitung dapat didekati dengan dua buah teknik pengumpulan data/cuplikan data. Pertama adalah bila kajian memerlukan keseluruhan populasi, sehingga tidak satupun yang tercecer dari aktivitas ini, yakni di perlukan pendekatan sensus. Sensus berupaya untuk mencacah keseluruhan populasi. Contohnya adalah sensus penduduk yang dilakukan oleh BPS (Badan Pusat Statistik) Republik Indonesia. Sementara jika menggunakan sebagian tertentu dari populasi, maka peneliti telah melakukan sampling atas populasi tersebut, sehingga riset yang dilakukan adalah survey.

4.2 Kriteria & Proses Penyusunan Kerangka Sampel

Salah hal yang elementer dalam proses penelitian adalah mendisain kerangka atau *frame of sample*, sebagai instrumen yang akan menuntun peneliti untuk mengambil data dari sampel yang telah ditentukan. Hal ini menjadi penting mengingat sampel harus representatif, mampu mewakili realitas yang dipancarkan populasi dengan sebaik-baiknya. Sampel adalah refleksi langsung dari populasi, dimana potret realitas yang akan di data berada sepenuhnya dalam sampel tersebut. Kerangka diperlukan agar dalam kerjanya, peneliti memiliki arah dan tujuan yang pasti.

Sebagaimana dipahami bersama, bahwa sampel merupakan cuplikan/sebagian dari populasi dimana karakteristik dan realitas yang berada padanya akan di cermati. Secara jumlah, sampel lebih sedikit atau sama dnegan jumph populasi. Realitas dalam objek yang akan diteliti yang berada di dalam sampel disebut sebagai elemen/unit sampel. Unit sampel inilah yang diambil dari konstruksi kerangka sampel.

Dalam pratik riset, peneliti sangat berperan penting didalam menentukan kerangka sampling ini, terutama dikaitkan dengan tujuan penelitian itu sendiri. Penentuan unit sampel dan unit analisis pun, sangat erat atau ditentukan oleh bagaimana tujuan penelitian itu sendiri. Unit sampel dalam riset objeknya bisa sama persis dengan unit analisisnya, namun bisa juga berlainan. Ilustrasinya adalah, jika peneliti mengambil program studi universitas X sebagai unit sampel, maka unit analisisnya bisa para mahasiswa dalam program studi tersebut. Dalam kasus itu unit sampel berbeda dengan unit analisis. Berbeda misalnya dengan peneliti yang hendak mencermati tingkat disiplin mahasiswa universitas Y , maka dapat diambil secara random populasi yakni mahasiswa yang kuliah di universitas Y tersebut. Disinilah terlihat bahwa unit sampel sama dengan unit analisis, yakni mahasiswa yang kuliah di universitas Y yang kebetulan terpilih sebagai responden.

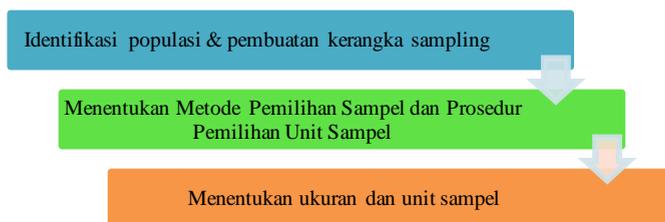
Satu hal penting yang wajib peneliti sadari adaah relasi sampel dan populasi harus intim, dalam arti sampel harus merefleksikan populasi dan menjadi representasi yang benar dari realitas di populasi. Dengan demikian di dalam pengambilan sampel wajib mencermati karakter sampel itu sendiri, yang antara lain adalah: (1) Sampel bisa menggambarkan secara meyakinkan dari populasi yang akan diteliti, (2) Sampel dapat menentukan tingkat presisi sebuah riset, yakni menentukan penyimpangan standar dari taksiran yang

diperoleh, (3) Sampel bersifat sederhana, bisa memberi penjelasan komprehensif, membantu mereduksi waktu, tenaga dan finansial penelitian.

Selanjutnya, bagaimanapun akurasi sebuah riset juga di tentukan oleh sejauhmana sampel itu akurat. Akurasi sampel menjadi pintu masuk penting dalam menentukan akurasi pada proses analisis data nanti. Terdapat beberapa cara untuk meningkatkan derajat akurasi dari sebuah sampel, yang diantaranya adalah:

- a. Pemilihan sampel sebaiknya dilandasi oleh proksi yang benar – benar tepat dan sesuai tujuan riset yang ditentukan. Misalnya kajian tentang kualitas komunikasi organisasi pada birokrasi pemerintahan dan organisasi swasta. Dalam hal ini sampel dibagi menjadi 2 bagian, birokrasi pemerintahan dan organisasi swasta. Ada beberapa proksi untuk mengukur kualitas komunikasi organisasi, dan dapat dipilih misalnya gaya kepemimpinan. Apabila proksi gaya kepemimpinan tidak bisa untuk mengukur kualitas komunikasi organisasi pada dua jenis perusahaan tersebut, maka proksi itu tidak akurat,
- b. Tidak menggunakan sampel yang bersifat bias atau semu. Pembiasan data oleh karena sampel yang keliru ini perlu diantisipasi dengan seksama, sehingga kajian menjadi akurat. Contoh pembiasan sampel misalnya, ketika kita ingin melihat kualitas komunikasi organisasi di organisasi pemerintahan, misalnya kementerian X, maka sampel sebaiknya mulai birokrasi kementerian pusat hingga tingkat kanwil (propinsi) dan tingkat cabang (kabupaten). Bila sampel tidak menjangkau itu, maka tidak akan akurat.

Untuk mendapatkan sampel yang akurat, terdapat proses yang mesti dilalui oleh peneliti dengan menempuh tahapan dan sistematika tertentu. Adapun tahapan yang sederhana dapat kita lakukan sebagai berikut:



Gambar : Tahapan Menentukan Sampel

Sumber : Olahan Penulis

Dari gambar diatas dapat dijelaskan :

- i. Mengidentifikasi populasi dapat dilakukan menentukan dengan jelas populasi apa yang diteliti, yang dikatakan sebagai populasi sasaran /*target population*. Misalnya akan diteliti hubungan jiwa kewirausahaan dengan kinerja industri UMKM di kota Medan., maka target populasinya adalah pelaku UMKM di kota Medan tersebut. Berikutnya dilakukan penentuan kerangka sampling, yakni daftar elemen populasi yang merupakan dasar dari penentuan sampel. Misalnya, populasi target adalah warga kota Medan hasil sensus penduduk tahun 2020, andai peneliti menggunakan daftar dari BPS, maka ada kemungkinan data tersebut belum memuar warga Medan yang tinggal mulai Januari 2021 misalnya. Maka orang tersebut meskipun sudah ber KTP Medan, namun belum terdata oleh sensus penduduk. Lain halnya jika peneliti menggunakan data populasi sasaran dari data Kantor Dukcapil.
Ilustrasi diatas memberikan pemahaman bagi kita bahwa dalam merancang kerangka sampel, kriteria yang sebaiknya dipatuhi adalah sebagai berikut: (1) Lengkap namun ada limitasi. Lengkap adalah sebaiknya tidak ada data yang tercecer atau ketinggalan. Ada limitasi dalam arti hanya hal yang kaitannya dengan unsur-unsur sampel saja yang dituliskan, (2) Tidak tumpang tindih dan faktual. Unsur – unsur sampel harus jelas, tidak tumpang tindih, di tulis secara berlebihan dan sebagainya. Faktual artinya berdasar fakta empiris, yang juga bermakna dapat ditelusuri dengan mudah di lapangan.
- ii. Selanjutnya, ada beberapa panduan yang bisa dimanfaatkan oleh peneliti guna menentukan metode pengambilan sampel yang akan ditempuh yakni sebagai berikut: (1) Sampling dengan Pengembalian/*Sampling with Replacement*. Teknik ini mirip arisan dimana orang terpilih tidak mau menerima hasilnya lalu di kembalikan. Dalam teknik sampel model ini, satuan sampling terpilih dikembalikan ke populasi sebelum pemilihan berikutnya, (2) Sampling tanpa Pengembalian/*Sampling without Replacement*. Dalam model ini satuan sampling yang telah terpilih tidak dikembalikan lagi ke dalam populasi. Dalam praktik penelitian, cara sampling seperti ini cukup sering digunakan. Tipe ini ada dua jenis, yakni: (a) Sampling Peluang (*Probability Sampling*), dimana dalam pemilihannya menyangkut probabilitas, dengan keterpilihan

satuan sampling dapat diketahui, yang berikutnya (b) Sampling Non peluang (*Nonprobability Sampling*) yang tidak melibatkan unsur probabilitas, sehingga dapat dilakukan langsung secara acak dan sederhana.

Tabel Jenis Sampling Pada Probability Sampling

Jenis-Jenis Sampling		Deskripsi
Simple Sampling	Random	Setiap elemen populasi berpeluang sama dipilih sebagai sampel
Systematic Sampling		Memilih dan memilih sampel dari populasi dari awal dan pemilihan sampel didasarkan pada urutan elemen.
Stratified Sampling	Random	Pembagian populasi menjadi kelompok dan dengan acak memilih sub-sampel dari masing-masing kelompok
Cluster Sampling	Random	Kumpulan yang heterogen diidentifikasi lebih awal, baru dipilih dengan acak, dan semua hasil pengacakan itu diteliti.
Multistage Sampling	Area	Peneliti memilih domain mikro pada tiap tahapannya, lalu mengombinasikan keempat teknik sampel itu.

Tabel Jenis Sampling Pada Nonprobability Sampling

Jenis-Jenis Sampling		Deskripsi
Convenience Sampling		Teknik pengambilan sampel paling sederhana atau ekonomis.
Accidental Sampling		Teknik pengambilan sampel dengan satuan sampling didapat dengan sembarang
Purposive Sampling/ Judgement Sampling		Teknik pengambilan sampel dengan pelibatan pakar / ahli bidang sedang diteliti.
Quota Sampling		Teknik pengambilan sampel melalui klasifikasi populasi, menentukan proporsi sampel tiap kelas, menetapkan kuota untuk interviewer.
Snowball Sampling		Responden awal dipilih via probabilitas, responden berikutnya atas rekomendasi responden sebelumnya tersebut.

Didalam riset – riset sosial, hal yang sering dipermasalahkan besarnya sampel, misalnya melibatkan minimal 10% dari besaran populasi yang ada. Sebenarnya ukuran sampel bisa bervariasi sebagaimana variasi yang ditunjukkan oleh populasinya, sehingga kian besar jumlah populasi, maka semakin besar juga ukuran sampelnya. Hal ini untuk memenuhi supaya perkiraan atas parameter populasi dapat akurat. Berkaitan dengan konteks ini,

ada faktor-faktor yang perlu diperhatikan ketika menentukan besaran sampel:

- (a) *Tingkat homogenitas populasi*. Semakin homogen/seragam populasi semakin sedikit sampel yang dibutuhkan, namun semakin tidak seragam/heterogen suatu populasi semakin besar sampel yang dibutuhkan untuk memastikan akurasi,
- (b) *Jenis sampling yang digunakan*. Riset yang digunakan untuk menguji parameter diperlukan jumlah sampel lebih besar;
- (c) *Kedalaman analisis riset*. Semakin mendalam analisis yang diinginkan sampel makin besar;
- (d) *Skala pengukuran* yakni nominal, ordinal, ratio dan interval;
- (e) Besaran *bound of error*, derajat kepercayaan (*level of significance*) atau α , kuasa uji (*power of test*) atau β yang dikehendaki.

4.3 Menentukan Ukuran Besaran Sampel

Sampel dalam penelitian, sampel ditentukan besarnya menurut kebutuhan yang ada. Secara matematis dapat dirumuskan beberapa perhitungan yang representatif untuk menentukan ukuran sampel tersebut.

a) Ukuran sampel untuk proporsi

Jika umpamanya peneliti ingin memperkirakan proporsi sebuah populasi, maka jumlah sampel diambil secara acak adalah :

$$n = \left[\frac{z_{(1-\alpha/2)} \sqrt{\pi(1-\pi)}}{\delta} \right]^2$$

Catatan :

π = perkiraan persentase /proporsi kejadian.

δ = *bound of error*.

Pada rumus diatas terdapat unsur parameter yang tidak diketahui, yaitu π . Jika tak ada informasi mengenai nilai π , maka dapat anggap bahwa harga $\pi = 0,5$, sehingga rumus diatas menjadi:

$$n = \left[\frac{z_{(1-\alpha/2)}}{2\delta} \right]^2$$

Seorang direktur marketing sebuah perusahaan makanan instan, ingin mendapatkan keterangan objektif mengenai keputusannya meluncurkan varian baru produk makanannya. Jika hasil penelitian harus mempunyai derajat kepercayaan 95 % dan *bound of error* $\delta = 5\%$, berapa orang yang harus diambil sebagai sampel agar hasil penelitian memenuhi ketentuan di atas.

Dengan tingkat kepercayaan 95% atau $\alpha = 5\%$ diperoleh harga:

$Z_{tabel} = Z_{(1-\alpha/2)} = Z_{1,96}$. Diketahui *bound of error* $\alpha = 5\% = 0,05$.

$$\text{Jadi } n = \left[\frac{Z_{(1-\alpha/2)}}{2\delta} \right]^2 = \left[\frac{1,96}{(2)(0,05)} \right]^2 = 385$$

Mengikuti rumusan diatas, jumlah sampel yang harus diambil sebagai sumber data dalam penelitian itu paling tidak sejumlah paling tidak 385 responden.

b) Ukuran Sampel untuk Rata-Rata

Jika peneliti ingin menaksir rata-rata populasi, maka sampel yang harus diambil acak yakni :

$$n = \left[\frac{Z_{(1-\alpha/2)} \sigma}{\varepsilon} \right]^2$$

dengan:

σ = adalah simpangan baku populasi untuk variabel yang sedang diteliti

ε = adalah bound of error

Secara khusus dapat dikatakan bahwa rumus tersebut mengandung parameter yang sulit sekali diketahui. Dalam level pragmatis, penjelasan tentang nilai simpangan baku populasi σ ini dapat diperoleh dengan beberapa cara yakni: (1) Melalui hasil penelitian terdahulu yang serupa; (2) Pendapat para pakar pada bidang yang diteliti; dan (3) Penelitian pendahuluan (pilot survey). Andai simpangan baku populasi σ tidak diketahui, maka ukuran sampel dapat dihitung dengan *iterative method*, misalnya adalah :

Taruh kata seorang dosen ingin meneliti persepsi masyarakat propinsi Sumatera Utara terhadap program pemerintah yakni PKH (Program Keluarga Harapan). Jika dalam riset akan di perkirakan rata-rata skor persepsi / sikap masyarakat terhadap PKH dengan ketentuan hasil penelitian memiliki tingkat kepercayaan 95% dan *bound of error* untuk rata-rata skor adalah 4, maka dapat dihitung ukuran sampel minimal sebagaimana diuraikan dibawah ini Penelitian itu menggunakan parameter yang akan diduga adalah rata-rata skor persepsi masyarakat Provinsi Sumatera Utara dengan *simple random sampling*. Tingkat kepercayaan yang dikehendaki adalah 95% atau $\alpha = 5\%$ dan *bound of error* $\varepsilon = 4$. Tampak tidak ada data soal simpangan baku populasi, sehingga ukuran sampel dapat dikalkulasi dengan metode iteratif (*iterative method*). Di misalkan diambil sampel acak dengan jumlah 8, maka dapat dituliskan $n = 8$ orang, yang kemudian dilakukan interview untuk

mengetahui persepsi mereka terhadap PKH. Data yang di dapat adalah 127, 136, 145, 129, 125, 137, 134, 130.

Dapat dihitung kemudian adalah simpangan baku $s = \sqrt{\frac{\sum x_1^2 - \frac{(\sum x_1)^2}{n}}{n-1}} = 5,46$

$$\text{Iterasi pertama } n_1 = \left[\frac{z_{(1-\alpha/2)} s}{\varepsilon} \right]^2 = \left[\frac{z_{0,975}^{(5,46)}}{4} \right]^2 = \left[\frac{(1,96)(5,46)}{4} \right]^2 = 7,16 = 8$$

$$\text{Iterasi kedua } n_2 = \left[\frac{t_{(1-\alpha/2; n_1-1)} s}{\varepsilon} \right]^2 = \left[\frac{t_{0,975; 8-1}^{(5,46)}}{4} \right]^2 = \left[\frac{(2,365)(5,46)}{4} \right]^2 = 10,43 = 11$$

Tampak hasil iterasi 1 dan iterasi 2 berbeda, sehingga dilakukan iterasi 3 yakni:

$$n_3 = \left[\frac{t_{(1-\alpha/2; n_2-1)} s}{\varepsilon} \right]^2 = \left[\frac{t_{0,975; 11-1}^{(5,46)}}{4} \right]^2 = \left[\frac{(2,228)(5,46)}{4} \right]^2 = 9,26 = 10$$

Iterasi 2 dan 3 masih berbeda, maka dilakuka iterasi ke 4 sebagai berikut :

$$n_4 = \left[\frac{t_{(1-\alpha/2; n_3-1)} s}{\varepsilon} \right]^2 = \left[\frac{t_{0,975; 10-1}^{(5,46)}}{4} \right]^2 = \left[\frac{(2,262)(5,46)}{4} \right]^2 = 9,54 = 10$$

Terlihat bahwa itersi 4 sama dengan iterasi 3, maka kalkulasi tuntas pada iterasi 4. Hasil perhitungan iterasi 4 menunjukkan $n = 10$, sehingga $n = 8$, wajib ditambah dengan 2 orang responden lagi, dan inilah yang di sebut dengan *Freund method*.

SOAL-SOAL LATIHAN

1. Apa yang dimaksudkan dengan populasi dan sampel dalam penelitian ilmiah? Jelaskan disertai contoh.
2. Bagaimana kriteria penentuan sampel dalam penelitian? Jelaskan disertai contoh.
3. Bagaimana prosedur penentuan kerangka sampel penelitian ilmiah?
4. Akurasi sebuah riset ilmiah, salah satunya ditentukan oleh penentuan sampel yang juga akurat. Bagaimana syarat agar pengambilan sampel yang akurat? Jelaskan.
5. Seorang direktur marketing sebuah perusahaan makanan instan, ingin mendapatkan keterangan objektif mengenai keputusannya meluncurkan varian baru produk makanannya. Jika hasil penelitian harus mempunyai derajat kepercayaan 95 % dan bound of error $\delta = 5\%$, berapa orang yang harus diambil sebagai sampel agar hasil penelitian memenuhi ketentuan di atas.

BAB V

PENTAHAPAN PENGUJIAN DENGAN STATISTIKA DALAM RISET

5.1 Pendugaan Dalam Penelitian

Dalam dunia penelitian, permasalahan yang hadir mesti di jawab dengan melaksanakan penelitian itu sendiri hingga selesai. Namun dalam metodologi riset, sesungguhnya jawaban itu dapat diduga diawal penelitian melalui pendugaan jawaban, atau hipotesis. Dengan kata lain, hipotesis merupakan jawaban yang sifatnya masih teoretis atas pertanyaan penelitian. Pendugaan jawaban atau hipotesis ini, di konstruksi melalui teori – teori yang digunakan. Ketika hipotesis telah dirumuskan sebagai gambaran jawaban pertanyaan, kemudian peneliti melakukan penelitian dengan prosedur ilmiah sehingga didapatkan informasi yang bisa menjawab, apakah dugaan awal tadi terbukti benar atau tidak. Didalam corak riset relasional, hipotesis juga telusur dengan melihat hubungan antara variabel – variabel yang telah ditentukan. Demikian pula dengan penentuan jenis uji statistik yang akan digunakan, juga didasarkan pada asumsi – asumsi yang telah dirumuskan, apakah dengan uji parametrik atau non parametrik.

Dengan menyusun dugaan sementara atau dugaan awal maka sebenarnya peneliti telah melakukan beberapa hal yang penting yakni jalan yang memudahkan atau mengikuti kaidah ilmiah penelitian. Dengan hipotesis peneliti akan: (1) Mempunyai tuntunan/pedoman dalam melakukan penelitian lebih lanjut, (2) Mendapatkan arah-tujuan penelitian menyangkut variabel, ukuran dan lainnya, (3) Mendapatkan imajinasi logis untuk mendisain penelitian hingga rancangan kesimpulan dari riset yang dilakukan.

Selanjutnya, seorang peneliti dalam metode kuantitatif harus merumuskan hipotesis, dan bagaimana hipotesis penelitian tersebut harus di uji nanti. Pentahapan uji hipotesis secara statistika dilakukan sebagai berikut: (1) Peneliti menyusun formulasi hipotesis nihil (H_0) dan hipotesis alternatifnya (H_1) atau hipotesis kerja/operasional, (2) Peneliti memilih satu model uji statistika yang kompatibel, dalam arti yan corak-modelnya paling mendekati asumsi, (3) menentukan taraf signifikansi (α) dan besarnya sampel penelitian (n), (4). Menentukan distribusi sampling, arah uji, daerah

kritis dan kearakteristik pengujian H_0 , (5) Menghitung nilai uji statistiknya dengan data yang telah didapat, (5). Peneliti lalu mengambil keputusan atau kesimpulan pengujian, yaitu apakah H_0 diterima atau H_0 ditolak atau menerima H_1 , menurut taraf signifikansi yang dipilih.

5.2. Pendugaan/Hipotesis Nihil, Hipotesis Kerja dan Pemilihan Uji Statistik

Permulaan dalam sistem dan prosedur pengujian adalah merumuskan dugaan nihil atau hipotesis nihil (H_0) dan hipotesis kerja (H_1). Secara ideal normative H_0 dinyatakan dalam kalimat “Tidak ada perbedaan ($\mu_1 = \mu_2$)”, “tidak ada pengaruh ($\beta = 0$)”, atau “tidak ada korelasi ($\rho = 0$)”. Hipotesis nihil ini bersifat teoretis, yang harus dibandingkan dengan hipotesis kerja / H_1 yang telah dilakukan operasionalisasi terhadapnya, melalui logika deduksi dan *derivate* teori dan kemudian di uji secara empirik.

Sebagai ilustrasi, dalam perspektif ekonomi para pelaku usaha yang bisa melakukan inovasi produk dan pemasaran akan memenangi persaingan pasar, di bandingkan dengan pelaku usaha yang tidak melakukan inovasi produk dan pemasaran. Kalimat diatas merupakan bahan – bahan untuk menyusun hipotesis penelitian. Dalam hal ini peneliti telah melakukan prediksi mengenai kecenderungan yang terjadi atas fenomena di atas. Hipotesis riset itu dinyatakan dalam bentuk operasional, yaitu berupa hipotesis kerja (H_1), sehingga H_1 akan berbunyi $\mu_1 \neq \mu_2$, yang maknanya adalah bahwa tingkat inovasi produk dan pemasaran rata-rata dari populasi pelaku usaha tidak sama. Disisi lain H_0 akan berbunyi $\mu_1 = \mu_2$, yang artinya bahwa derajat inovasi produk dan pemasaran para pelaku usaha tersebut sama atau tak beda. Selajutnya andai data empiris menampilkan bahwa H_0 dapat ditolak, maka H_1 dapat diterima dan ini mendukung hipotesis penelitian serta teori yang mendasarinya. Selanjutnya, jika hipotesis riset hanya menyatakan, bahwa kedua kelompok populasi itu berbeda angka rata-rata (meannya), maka H_1 adalah $\mu_1 \neq \mu_2$. Namun, andai teorinya memprediksikan arah perbedaan itu, yakni bahwa kelompok tertentu memiliki mean yang lebih besar daripada kelompok lain, maka H_1 berbunyi $\mu_1 > \mu_2$ atau jika sebaliknya $\mu_1 < \mu_2$ (di mana $>$ berarti lebih besar dan $<$ berarti lebih kecil daripada). Kemudian, peneliti membutuhkan landasan yang *make sense* atau rasional, memilih jenis uji

statistik tersebut. Pemilihan didasarkan kepada dugaan atau asumsi dan syarat pengukuran variabel-variabel risetnya.

5.3 Menentukan Taraf Signifikansi dan Besaran Sampel Penelitian

Jika hipotesis telah ditentukan dan bentuk uji statistiknya juga sudah ditentukan, maka hal berikutnya yang ditempuh adalah menentukan taraf signifikansi (α) dan jumlah sampel (n). Pra data real/lapangan dikoleksi, peneliti menetapkan semua sampel, atau himpunan sampel dari domain atau populasi yang muncul andai H_0 benar. Kemudian dikonstruksi himpunan bagian dari sampel yang mungkin, sehingga andai H_0 benar, meskipun probabilitasnya lemah bahwa sampel yang peneliti dapati itu masuk dalam himpunan bagian itu. Bila kemudian dalam riset peneliti mengamati sampael termasuk dalam himpunan bagian tersebut, maka H_0 ditolak.

Seorang peneliti sampai pada ketetapan hati untuk menolak H_0 dan menerima H_1 , adalah ketika uji statistik menghasilkan harga yang barangkali munculnya di bawah $H_0 = /<$ (sama dengan atau lebih kecil) dari suatu harga alpha (α). Nilai atau harga kemungkinan kecil inilah yang disebut dengan taraf signifikansi (*the level of significance*). Disini dapat dimaknai betapa taraf signifikansi adalah probabilitas tertolaknya H_0 apabila realitasnya H_0 benar. Lazimnya, riset – riset menggunakan taraf signifikansi 0,05 dan 0,01.

Kesimpulan tentang H_0 , berpeluang terjadi kesalahan, yakni kesalahan tipe I (*type I error*), terjadi ketika peneliti menolak H_0 , sementara realitasnya H_0 itu benar. Selanjutnya, ada kesalahan tipe II (*type II error*), terjadi ketika peneliti menerima H_0 meski realitasnya H_0 itu salah. Kemungkinan terjadi kesalahan tipe I ditandai dengan α , dimana semakin besar nilai atau harga α , akan semakin besar juga probabilitas H_0 ditolak secara salah. Umumnya probabilitas melakukan kesalahan tipe I lebih besar. Kesalahan tipe II diberi symbol β , menunjukkan probabilitas melakukan kesalahan tipe II.

Kesimpulan Penelitian	Kenyataan	
	H_0 Benar	H_0 Salah
Terima H_0	Keputusan tepat ($1 - \alpha$)	II Kesalahan tipe (β)
Tolak H_0	Keputusan tipe I (α)	Keputusan tepat ($1 - \beta$)

Riset yang ideal yang memenuhi kaidah ilmiah biasanya menetapkan besaran angka α dan β terlebih dahulu, sebelum riset tersebut dilaksanakan. Sebab penentuan besaran angka α dan β akan menjadi dasar dalam menentukan besar sampel (n) yang wajib diambil peneliti, guna menjalankan risetnya dengan panduan uji statistik terpilih. Jika besaran α dan n telah dipastikan pada akhirnya nilai β sudah ada dan bernilai tertentu. Besaran α dan β berhubungan secara bolak balik dimana deskresi pada nilai α akan meningkatkan nilai β untuk besara n tertentu. Namun andai, seorang peneliti berniat mereduksi probabilitas melakukan kesalahan, maka n (besar sampel) harus diperbesar.

5.4 Distribusi Sampling Dan Nilai Uji Statistik

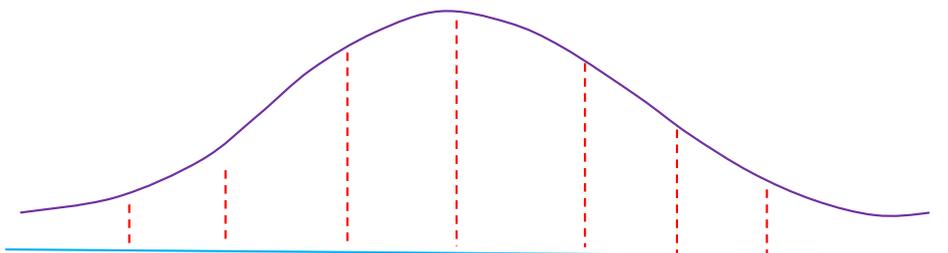
Ketika seorang peneliti telah memasuki tahapan sudah menentukan jenis uji statistik yang akan digunakan untuk menganalisis data empiris guna menyelesaikan permasalahan yang telah ditentukan dalam kajian itu, maka peneliti sampai pada bagaimana melakukan distribusi sampling dari uji statistik yang dipilihnya tersebut. Per definisi dapat dikatakan bahwa distribusi sampling adalah distribusi teoretis yang didapatkan peneliti saat mengambil seluruh sampel yang mungkin dan dengan besaran/ukuran (n) yang sama. Pengambilan itu dari sebuah domain atau populasi dengan besaran mean μ dan deviasi standar σ , serta pengambilannya dilakukan dengan sembarang.

Dalam pemahaman lain, distribusi sampling merupakan distribusi di bawah harga H_0 dari keseluruhan harga yang mungkin dilalui oleh ukuran – ukuran tertentu jika harga statistik dikalkulasi dari cuplikan/sampel yang setara besarnya dengan pengambilan sembarang/acak. Dalam praktik statistika, distribusi sampling suatu nilai statistik lazimnya akan memenuhi

suatu teori batas memusat atau teori central limit (*central limit theorem*) dengan kaidahnya sebagai berikut :

Apa bila sebuah variabel penelitian berdistribusi normal dengan rata-rata/mean = μ dan deviasi standar = σ , dan diambil dari sampel acak berukuran (n), maka himpunan rata – rata/mean dari sampel-sampel tersebut (\bar{X}) akan berdistribusi kurang lebih normal, dengan *mean* dari himpunan *mean* adalah μ dan simpangan baku atau deviasi si standar sebesar $\sigma\sqrt{n}$. Dengan demikian andaikan ukuran sampel relative besar ($n > 30$), maka distribusi sampling rata rata (\bar{X}) akan menampilkan karakteristik :

1. Mengikuti realitas distribusi yang normal;
2. Memiliki harga *mean* yang sama dengan harga *mean* populasinya ($\mu_x = \mu$);
3. Memiliki simpangan baku atau deviasi standar sama dengan deviasi standar populasi dibagi oleh akar nilai sampel ($\sigma_{\bar{x}} = \sigma\sqrt{n}$).



X $-3\sigma_x$ $-2\sigma_x$ $-\sigma_x$ U_x $+\sigma_x$ $+3\sigma_x$ $+3\sigma_x$

Selanjutnya, hipotesis kerja / alternatif yang memberi pernyataan “ada perbedaan”, misalnya $H_1 : \mu_1 \neq \mu_2$ akan diperlakukan dengan pengujian dua pihak. Bila hipotesis kerja/alternatif yang menyatakan “lebih besar dari”, misalnya $H_1 : \mu_1 > \mu_2$ diuji satu sisi kanan atau *one-tailed test* atau *one-sided test-right*. Sebaliknya, untuk hipotesis kerja/alternatif yang menyatakan “lebih kecil dari”, misalnya $H_1 : \mu_1 < \mu_2$ dilakukan uji satu sisi kiri atau *left sided test*.

Keyakinan dalam pengambilan keputusan berkenaan dengan hipotesis teoretik atau H_0 berasal dari data empirik sampel kemudian perlu dihitung harga uji statistiknya. Harga ini merupakan derajat penyimpangan nilai statistika sampel (contoh \bar{X}) dari besaran parameter yang diduga (contoh

μ_0) dalam satuan simpangan baku distribusi sampling nilai statistiknya (contoh $\sigma \sqrt{n}$). Dengan demikian, uji statistiknya dapat dirumuskan adalah:

$$Z = \frac{\bar{X} - \mu_0}{\sigma \sqrt{n}}$$

di mana:

\bar{X} = adalah nilai rata-rata sampel

μ_0 = adalah nilai rata-rata awal

σ = adalah nilai simpangan baku

n = adalah jumlah sampel

Pengujian statistik itu berlaku untuk uji rataan populasi dan untuk sampel tunggal dengan ukuran relatif besar dengan variabel yang menggunakan skala interval dan skala rasio. Dengan demikian variasi nilai yang dihasilkan akan sangat ditentukan oleh pemilihan cara uji statistik oleh peneliti itu sendiri.

5.5 Proses Pengambilan Keputusan Berdasar Hasil Pengujian

Sesi ujung dalam proses ilmiah menguji hipotesis adalah tindakan penyimpulan terhadap hasil apakah H_0 diterima atau H_0 ditolak, yang didasarkan pada derajat signifikansi yang telah ditentukan. Dari situ dapat digambarkan bahwa andaikan peluang suatu harga tertentu dalam distribusi sampling di bawah H_0 sangat kecil, maka ada dua kemungkinan yakni (1) hipotesis nihilnya salah, atau (2) terjadinya peristiwa yang langka dan kecil sekali kemungkinan terjadinya. Kemungkinan itu dinyatakan sebagai α , yang menunjukkan probabilitas menolak H_0 padahal kenyataannya H_0 benar atau yang disebut dengan kesalahan tipe I.

Bila probabilitas tentang harga teramati dari hasil sebuah pengujian adalah lebih kecil dari α yang sudah ditentukan sebelumnya, kesimpulannya adalah H_0 ditolak, inilah yang disebut dengan nilai yang “signifikan”. Secara teoretik dapat dirumuskan bahwa suatu nilai/ harga dikatakan sebagai “signifikan” adalah ketika tingkat kemungkinan yang diperoleh atau p -value $< \alpha$. Bahasa lainnya adalah kesalahan yang terjadi (p -value) lebih kecil dari tingkatan toleransi kesalahan yang telah ditetapkan sebelumnya (α).

Ilustrasi Permasalahan :

Seorang kepala dinas peternakan di daerah X mengatakan bahwa selama ini produksi susu sapi murni hanya 10 liter per hari per sapi. Hal ini menurutnya sudah cukup melebihi kapasitas sapi dalam memproduksi air susu. Seorang mahasiswa ingin meneliti hal tersebut untuk mendapatkan kepastian kapasitas produksi peternakan sapi perah tersebut. Dia lalu menyusun hipotesis kerja yang mengatakan bahwa “kemampuan produksi rata – rata sapi perah di daerah X sama dengan 10 liter per hari per sapi”. Kemudian di tarik sampel dengan sembarang atau random sebanyak 64 dari para peternak yang ada di daerah X tersebut. Dari pengukurannya didapatkan data rata – rata hasil susu dari sampel tersebut sebanyak 11 liter per kambing per hari. Menurut *central limit theorema* distribusi sampling (\bar{X}) dari sampel berukuran 64 dari populasi berdistribusi normal, dengan *mean* sama dengan *mean* populasinya dan deviasi standar sama dengan deviasi standar populasi dibagi akar anggota sampel. Andai diketahui bahwa simpangan baku populasinya 4 liter, dan ditetapkan taraf signifikansi 0,05.

Penyelesaian:

Dari data diatas, hipotesis dapat di uji sebagai berikut :

- a. Hipotesis statistiknya

H_0 : $\mu = 1$ liter; artinya produksi susu rata-rata per ekor per hari dari para peternak di Kabupaten X adalah 1 liter.

H_0 : $\mu \neq 1$ liter; artinya produksi susu rata-rata per ekor per hari dari para peternak di Kabupaten X tidak 1 liter.

- b. Taraf nyata sebesar 5% ($\alpha = 0,05$).
- c. Uji statistik yang digunakan seperti pada rumus berikut ini.

$$Z = \frac{\bar{X} - \mu_0}{\sigma \sqrt{n}}$$

di mana:

$$\bar{X} = 11 \text{ liter}$$

$$\mu_0 = 10 \text{ liter}$$

$$\sigma = 4 \text{ liter}$$

$$n = 64$$

sehingga $z = 2$

- d. Menentukan titik kritis, perhatikan



Gambar : Titik kritis distribusi normal.

e. Karakteristik Pengujian

Pengujian mendapati diperoleh nilai z sebesar 2, dimana menurut distribusi normal nilai itu berdomain penolakan sehingga hipotesis permulaan yakni (H_0) ditolak. Selain itu jika di dibandingkan dengan tabel, Z - hitung ($Z = 2$) lebih besar dari z -tabel ($z = 1,96$) maka H_0 ditolak.

f. Kesimpulan:

Dengan demikian berdasarkan uji statistik yang berderajat kepercayaan 95% dapat dinyatakan, bahwa produksi susu rata-rata per ekor per hari dari para peternak di Kabupaten X adalah sama sekali tidak sama dengan 10 liter / lebih dari angka 10 liter.

SOAL – SOAL LATIHAN

1. Dengan menyusun dugaan sementara sebenarnya peneliti telah melakukan beberapa hal yang penting dan bermanfaat dalam menjalankan penelitiannya. Apa manfaat penyusunan hipotesis ?
2. Apa yang dimaksud hipotesis nihil dan hipotesis kerja ? Jelaskan disertai contoh !
3. Apa yang dimaksud teori batas memusat atau *central limit theorem*? Jelaskan!
4. Bagaimana yang dimaksud dengan distribusi sampling ? Jelaskan !
5. Bagaimana proses pengambilan keputusan berdasarkan hasil pengujian ? Jelaskan disertai contoh.

BAB VI

PENDUGAAN NILAI POPULASI DALAM STATISTIKA

Nilai populasi menjadi hal yang sangat elemeter dalam perhitungan penelitian, terutama dalam pendekatan kuantitatif. Besaran nilai populasi dapat diduga dengan menggunakan ilmu statistika, yakni melalui proses menentukan parameter. Pendugaan parameter dalam statistika dapat ditempuh dengan cara (1) *interval estimation*, dan (2) *point estimation*.

Dalam cara yang pertama yakni *interval estimation* atau pendugaan interval, dinyatakan bahwa penghitungan nilai statistik suatu sampel random dapat dilakukan menghitung dua batas nilai, di mana dengan derajat kemungkinan tertentu (*confidence interval*), nilai suatu parameter populasi yang akan diestimasi berada dalam batas nilai tersebut.

Sementara dalam cara kedua, yakni *point estimation* atau pendugaan titik dapat dilakukan dengan mengestimasi harga parameter populasi dengan satu nilai saja, yakni harga statistik sampelnya. Pada pendugaan titik ini, yang juga perlu diperhatikan adalah mengestimasi besarnya kesalahan yang dengan tingkat kemungkinan tertentu dialami pada saat menduga pendugaan harga parameter populasi itu. Mungkin sebagaimana dengan pengujian hipotesis estimasi nilai parameter populasi tersebut akan disandarkan pada teori probabilitas, dan juga distribusi sampling harga statistik dengan pendekatan “Central-limit theorem”. Persebaran atau distribusi sampling nilai statistik ini dapat berupa: (1) distribusi sampling nilai rata-rata (mean); (2) distribusi sampling nilai simpangan baku (deviasi standar); (3) distribusi sampling nilai variance; (4) distribusi sampling nilai proporsi; (5) distribusi sampling nilai beda dua rata-rata, dan (6) distribusi sampling nilai beda dua proporsi.

6.1 Pendugaan Harga Rata-Rata Pupulasi Untuk Sampel $n \geq 30$ dan $n < 30$

Dalam dunia statistika ada perbedaan perlakuan untuk besaran sampel-sampel dalam penelitian. Dalam konteks ini andai sebuah populasi mempunyai distribusi frekuensi normal dengan *mean* = μ dan simpangan baku = σ , dan ketika sampling tekniknya adalah random dari populasi

tersebut, maka himpunan mean (\bar{X}) dari sampel-sampel itu akan berdistribusi normal dengan rata-rata populasinya dan deviasi standar populasi dibagi dengan akar anggota sampel. Kemudian bila jika simpangan standar populasi tidak diketahui, maka standar *error* dari *mean* bisa diestimasi menggunakan simpangan standar sampel.

Penghitungan untuk sampel (baik n lebih dari atau sama dengan 30 / sampel besar, atau kurang dari 30/sampel kecil) yang tetapi populasinya berdistribusi normal dan σ diketahui, maka distribusi sampling harga *mean* dapat ditransformasikan menjadi distribusi normal standar dengan:

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \quad \text{atau} \quad \text{apabila sampel besar dan } \sigma \text{ tidak diketahui} \quad Z = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

berbagai nilai Z ini dapat diperiksa pada tabel kurva normal standar (lampiran A)

Estimasi harga *mean* populasi didasarkan pada taraf kepercayaan (*confidence level*) tertentu yang besarnya $1 - \alpha$. Bila digunakan *interval estimation* (pendugaan interval) rumusnya adalah:

$$\bar{X} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \text{atau}$$

$$\bar{X} - Z_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2} \frac{s}{\sqrt{n}} \quad \text{apabila } \sigma \text{ tidak diketahui}$$

Bila digunakan *point estimation* (pendugaan titik) maka μ diduga menggunakan \bar{X} dengan kemungkinan kesalahan $E - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ atau $E \leq + Z_{\alpha/2} \frac{s}{\sqrt{n}}$ apabila σ tidak diketahui.

$$\bar{X} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Bila σ tidak diketahui diganti s

$$\mu = \bar{X}$$

$$E \leq + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Bila σ tidak diketahui, diganti dengan s

Ilustrasi :

Seorang pengamat demografi ingin mengestimasi umur rata – rata anggota dewan perwakilan rakyat wanita di propinsi X. Langkah pertama yang ditempuh adalah dia mengambil dengan sembarang/acak sampel dari

populasi dimaksud dan mendapatkan jumlah $n = 120$ orang. Pengambilan sampel tersebut menghasilkan rata-rata umur para anggota dewan tersebut adalah 44,5 tahun. Kemudian simpangan standar = 9,75 tahun, dan derajat kepercayaan (*confidence level*) 0,95. Pertanyaannya, bagaimana perhitungan interval kepercayaan untuk usia rata-rata pada populasi anggota dewan perwakilan rakyat wanita di propinsi X?

Formulasi yang dapat untuk menghitung adalah :

$$\begin{aligned} \bar{X} - Z_{\alpha/2} \frac{s}{\sqrt{n}} &< \mu < \bar{X} + Z_{\alpha/2} \frac{s}{\sqrt{n}} \\ 44,5 - 1,96 \frac{9,75}{\sqrt{120}} &< \mu < 44,5 + 1,96 \frac{9,75}{\sqrt{120}} \\ 44,5 - 1,74 &< \mu < 44,5 + 1,74 \\ 42,76 &< \mu < 46,24 \end{aligned}$$

Dari perhitungan dan besaran – besaran yang di tentukan diatas, dapat ditarik simpulan sebagai berikut :

Pada *confidence level* 0,95, dapat di estimasikan bahwa usia rata-rata anggota dewan perwakilan rakyat wanita di propinsi X dalam interval antara 42,76 tahun hingga 46,24 tahun, yang dalam bahasa matematis adalah $P(42,76 < \mu < 46,24) = 0,95$. Namun jika dipandang dari estimasi titik maka rata-rata usia anggota dewan perwakilan rakyat wanita di propinsi X dihitung dengan pendugaan titik.

Apabila digunakan pendugaan titik, umur rata-rata para anggota dewan perwakilan rakyat wanita di propinsi X diperkirakan 44,5 tahun, pada probabilitas deviasi $E \leq 1,74$ tahun pada taraf kepercayaan 0,95.

Berdasarkan rumus E (error), kita dapat menentukan besar sampel (n) yang perlu ditarik agar diperoleh hasil dengan tingkat ketepatan tertentu. Tingkat ketepatan dikatakan dengan mengemukakan besarnya error maksimum yang mungkin dialami dalam melakukan pendugaan pada probabilitas tertentu.

$$\begin{aligned} E &= Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \\ n &= \left[\frac{Z_{\alpha/2} \cdot \sigma}{E} \right]^2 \end{aligned}$$

Sementara itu untuk memperhitungkan distribusi *harga mean* dengan jumlah sampel sembarang yang kecil ($n < 30$), pola distribusinya di luar

standar normal. Untuk sampel kecil ini mengikuti aturan atau rumus distribusi *student* dengan *mean* 0 dan distribusi simetris dengan derajat bebas (*Degree of freedom*) $n - 1$. Nilai standarnya dinyatakan dengan:

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

Rumus untuk pendugaan interval dinyatakan sebagai berikut:

$$\bar{x} - t_{\frac{\alpha}{2};(n-1)} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2};(n-1)} \cdot \frac{\sigma}{\sqrt{n}} \text{ atau } \bar{x} - t_{\frac{\alpha}{2};(n-1)} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2};(n-1)} \cdot \frac{s}{\sqrt{n}}$$

Berbagai nilai t dapat diperiksa pada tabel nilai *student t* (lampiran B).

Ilustrasi :

Dalam hal ini bisa dibandingkan atau dilihat kembali mengenai pendugaan harga *mean* untuk populasi yang besar atau ($n \geq 30$). Jika sampel acak wanita anggota dewan (tingkat kabupaten / kota) hanya hanya 25 orang ($n = 25$), dengan rata-rata usia $\bar{X} = 44,5$ tahun dan $S = 9,75$ tahun, maka:

$$\begin{aligned} \bar{x} - t_{0,025;24} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{0,025;24} \frac{s}{\sqrt{n}} \\ 44,5 - 2,064 \frac{9,75}{\sqrt{25}} < \mu < 44,5 + 2,064 \frac{9,75}{\sqrt{25}} \\ 44,5 - 4,025 < \mu < 44,5 + 4,025 \\ 40,475 < \mu < 48,525 \end{aligned}$$

Dengan demikian usia rata-rata wanita anggota dewan perwakilan rakyat tingkat II untuk propinsi Y diestimasi dalam rentang antara 40,475 tahun sampai 48,525 tahun (dibulatkan 40,5 tahun — 48,5 tahun).

6.2 Pendugaan Proporsi Pada Populasi

Pendugaan yang cukup presisi adalah dengan menggunakan \hat{p} ($\hat{p} = \frac{X}{n}$), dimana kenyataannya, domain populasi berukuran sangat besar, dan proporsi populasi tidak bergulir mendekati angka nol, maka anggapan saja tidak sangat mendekati nol atau satu maka distribusi sampling nilai proporsi bisa dianggap nyaris sama dengan distribusi normal, yakni ketika nilai *mean* sama dengan proporsi populasi serta penyimpangan standar.

$$\sigma_{\hat{p}} = \sqrt{\frac{\pi(1-\pi)}{n}} \text{ atau } s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Dalam rumusan yang baku dapat dituliskan distribusi sampling harga (\hat{p}) yang pada dasarnya adalah distribusi probabilitas normal, bisa ditransformasikan menjadi distribusi normal standart dengan:

$$z = \frac{\hat{p} - \pi}{\sigma_{\hat{p}}}$$

$$z = \frac{\hat{p} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \quad \text{atau}$$

$$z = \frac{\hat{p} - \pi}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \quad \text{karena } \pi \text{ tidak diketahui}$$

Jika ingin melakukan pendugaan proporsi sebuah populasi, yakni bila digunakan pendugaan interval, maka rumus yang dapat digunakan adalah :

$$\hat{p} - Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < \pi < \hat{p} + Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Selanjutnya, andaikata populasinya mempunyai normal, dan sampel acaknya berukuran besar ($n \geq 30$) maka distribusi sampling nilai penyimpangan standar (S) dianggap mendekati distribusi normal dengan *mean* sama dengan deviasi standar populasi ($\pi_s = \sigma$) dan penyimpangan standar sama dengan penyimpangan standar populasi dibagi oleh akar dari dua kali besar sampel. Nilai standarnya dinyatakan dengan:

$$Z = \frac{s - \sigma}{\sigma / \sqrt{2n}} \quad Z = \frac{s - \sigma}{\sigma_s}$$

Rumus estimasi harga σ dinyatakan dengan

$$S - Z_{\alpha/2} \frac{\alpha}{\sqrt{2n}} < \sigma < S + Z_{\alpha/2} \frac{\alpha}{\sqrt{2n}}$$

$$\frac{s}{1 + \frac{Z_{\alpha/2}}{\sqrt{2n}}} < \sigma < \frac{s}{1 - \frac{Z_{\alpha/2}}{\sqrt{2n}}}$$

Sebagai ilustrasi adalah: diketahui bahwa dari distribusi usia rata – rata anggota dewan perwakilan rakyat wanita di propinsi X, untuk $n = 120$ memiliki deviasi standar = 9,75 tahun. Pada tingkat kepercayaan 0,95, bagaimana perkiraan deviasi standar usia dari populasi anggota dewan perwakilan rakyat wanita pada propinsi X tersebut.

Dengan rumus tersebut diperoleh:

$$\frac{9,75}{1 + \frac{1,96}{\sqrt{2(120)}}} < \sigma < \frac{9,75}{1 - \frac{1,96}{\sqrt{2(120)}}}$$

$$\frac{9,75}{1+0,127} < \sigma < \frac{9,75}{1-0,127}$$

$$8,65 < \sigma < 11,17$$

6.3 Pendugaan Harga Beda Dua Mean Populasi dengan Sampel Besar dan Kecil

Pembahasan kali ini adalah pendugaan untuk dua mean dengan populasi sampel besar ($n_1 \geq 30$ dan $n_2 \geq 30$) dahulu. Andai kata dua domain populasi dengan distribusi normal mempunyai *mean* masing-masing sebagai μ_1 dan μ_2 , simpangan standar masing-masing σ_1 , σ_2 , . Selain itu sampelnya acak dengan ukuran $n_1 \geq 30$ dan $n_2 \geq 30$ untuk masing – masing populasi. Selanjutnya, himpunan harga beda dua *mean* ($\bar{x}_1 - \bar{x}_2$) memiliki distribusi normal, dengan *mean* sama dengan beda *dua mean* populasinya dengan deviasi standar :

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

Penghitungan harga standar bagi distribusi sampling harga *beda dua mean*, untuk sampel besar adalah :

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \text{atau}$$

$$Z = Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Kemudian, estimasi intervalnya dituangkan dengan rumusan sebagai berikut:

$$(\bar{x}_1 - \bar{x}_2) - Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Ilustrasi :

Konstruksi riset terhadap rata-rata umur anggota dewan perwakilan rakyat perempuan dimisalkan dilakukan di dua propinsi, yakni X dan Y. Di andaikan didapat informasi data empiris sebagai berikut :

$$n_1 = 120$$

$$\bar{X}_1 = 44,5 \text{ tahun}$$

$$S_1 = 9,75$$

$$n_2 = 100$$

$$\bar{X}_2 = 40,5 \text{ tahun}$$

$$S_2 = 8,5$$

Keterangan : angka indeks 1 menunjukkan propinsi X

Angka indeks 2 menunjukkan propinsi Y

Selanjutnya dengan derajat keyakinan 0,95, bagaimana estimasi beda umur rata-rata antara populasi anggota dewan perwakilan rakyat perempuan di dua provinsi X dan Y tersebut ?

$$(\bar{x}_1 - \bar{x}_2) - 1,96 \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + 1,96 \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$$

$$(44,5 - 40,5) - \sqrt{\frac{9,75^2}{120} + \frac{8,5^2}{100}} < (\mu_1 - \mu_2) < (44,5 - 40,5) + 1,96 \frac{9,75^2}{120} + \frac{8,5^2}{100}$$

$$1,59 < (\mu_1 - \mu_2) < 6,41$$

Kemudian, sebagaimana di jelaskan sebelumnya, perlakuan untuk sampel yang berharga kecil, yakni pada kisaran $n_1 < 30$ dan $n_2 < 30$, maka distribusi sampling nilai beda dua mean $\bar{x}_1 - \bar{x}_2$ mengikuti distribusi *student t*.

SOAL – SOAL LATIHAN

1. Bagaimana cara menduga besaran nilai populasi dalam ilmu statistika? Jelaskan!
2. Pengujian hipotesis estimasi nilai parameter populasi dapat berdasarkan pada teori probabilitas, dan juga distribusi sampling harga statistik. Dalam distribusi sampling terdapat pendekatan “Central-limit theorem”. Nah bagaimana persebaran atau distribusi sampling nilai statistik tersebut?
3. Dalam sebuah release, lembaga survey mengatakan telah melakukan pendugaan usia rata – rata anggota HIPMI (Himpunan Pengusaha Muda Indonesia) di propinsi X. Pihaknya telah mengambil sampel sejumlah 125 orang. Dari sampel itu diketahui usia rata – rata anggota HIPMI adalah 33,768 tahun. Jika standar deviasi 9,75 dan taraf kepercayaan 95 persen. Bagaimana gambaran interval usia rata rata anggota HIPMI di propinsi X tersebut?
4. Pendugaan terhadap interval usia rata rata anggota HIPMI yang sama dilakukan juga di kabupaten Z yang diperoleh data yakni, sampel tergolong kecil yakni hanya 30 orang. Sedangkan usia rata – rata adalah 33,768 tahun. Jika standar deviasi 9,75 dan taraf kepercayaan 0, 95. Bagaimana gambaran interval usia rata rata anggota HIPMI di kabupaten Z tersebut?
5. Survey pendugaan usia rata – rata anggota HIPMI juga dilakukan di propinsi Y. Rupanya lembaga survey ingin mengetahui perbedaan interval usia rata – rata anggota HIPMI di propinsi X dan Y. Dari upaya itu diketahui untuk propinsi X , sampel sebanyak 125 orang, usia rata–ratanya 33,768 tahun, standar deviasi 9,75 dan taraf kepercayaan 0,95. Sedangkan untuk propinsi Y diperoleh n sebesar 100 orang, dengan usia rata-rata 33,150 , standar deviasi 8,5 dan taraf kepercayaan 95 persen. Bagaimana pendugaan beda usia rata–rata antara kedua propinsi tersebut?

BAB VII

ANALISIS KORELASI

Dalam riset kuantitatif , utamanya riset – riset yang melibatkan dua atau lebih perubah atau variabel, ketika peneliti ingin membedah derajat keterkaitan hubungan atau relasi antar variabel tersebut maka diperlukan instrument analisis , yakni analisis korelasi. Jika di cermati, maka penampilan pola hubungan antar veriabel itu disebut dengan interaksi korelasi atau hubungan korelasi, yang dalam pengujian matematisnya membutuhkan tetapan atau koefisien relasi.

Angka tetapan korelasi linear yang positif menampilkan, bahwa kedua variabel berubah dalam arah yang sama. Dalam bahasa lain jika apabila variabel yang satu bertambah maka variabel yang lainnya juga bertambah. Disisi lain, pada koefisien korelasi linear yang negatif, andaikan variabel yang satu bertambah maka variabel lain berkurang atau dalam hal ini adalah berbanding terbalik. Selanjutnya, nilai r sama dengan nol dapat juga dimaknai bahwa kedua variabel tidak memiliki relasi linear, tetapi mungkin memiliki hubungan non-linear. Jika dilihat dari sudut pandang jumlah variabel, terdapat 3 macam analisis korelasi linear, yaitu: (1) Analisis Korelasi Linear Sederhana, (2) Analisis Korelasi Linear Parsial, (3) Analisis Korelasi Linear Berganda (*Multiple*).

Dalam bahasa statistika tetapan korelasi dinyatakan dengan simbol r , yakni bilangan yang menunjukkan derajat kuat-lemahnya hubungan antara dua variabel yang dimaksud. Tingkat besarnya tetapan korelasi dalam interval antara 0 sampai dengan 1 (untuk korelasi positif atau *direct correlation*) dan juga diantara 0 sampai dengan -1 (untuk korelasi negatif atau *inverse correlation*). Apabila nilai $r = 0$, maknanya adalah antara dua variabel tidak berkorelasi, apabila nilai $r = 1$ maknanya variabel berkorelasi positif secara sempurna, dan apabila jika harga $r = -1$ maknanya variabel berkorelasi negatif secara sempurna. Secara singkat harga tetapan korelasi (r) dalam interval antara -1 dan 1, atau $-1 \leq r \leq 1$. Kemudian untuk mempermudah pemahaman, berikut konstruksi penafsiran atas koefisien korelasi antar variabel yang di maksud.

No	Nilai Korelasi (r)	Penafsiran
1	0,00 – 0,19	Hubungan yang sangat kecil dan bisa diabaikan
2	0,20 – 0,39	Hubungan yang kecil (tidak erat)
3	0,40 – 0,69	Hubungan yang moderat
4	0,70 – 0,89	Hubungan yang erat
5	0,90 – 1,00	Hubungan yang sangat erat (sempurna)

Gambar : Penafsiran Tetapan Korelasi

7.1 Konstruksi Analisis Korelasi Linear Sederhana

Dalam level pragmatis, analisis korelasi sederhana dimanfaatkan peneliti untuk menjelaskan relasi atau hubungan antara dua buah variabel X dan Y. Dalam konteks ini, terdapat banyak koefisien korelasi yang cocok atau sesuai digunakan untuk menjelaskan hubungan antar variabel, namun semua akan tergantung pada jenis data dan skala pengukuran yang digunakan untuk menghimpun data empiris di lapangan. Sebagaimana telah di tulis pada sub bab sebelumnya mengenai skala pengukuran, maka di bawah ini akan di lihat penggunaannya yang dikaitkan dengan koefisien korelasi.

a. Data Nominal

Guna menjelaskan derajat hubungan korelasi antar variabel pada data yang diorganisasi dengan penskalaan nominal dapat digunakan koefisien korelasi ϕ (ϕ) yang formulasikan dengan:

$$\phi = \sqrt{\frac{X_p^2}{n}}$$

$$\text{Dimana } X_p^2 = \sum_i \sum_j \frac{(e_{ij} - o_{ij})^2}{e_{ij}}$$

e adalah frekuensi harapan (*expected*), dan

o adalah frekuensi amatan (*observed*).

b. Data Ordinal

Tetapan korelasi yang digunakan untuk menjelaskan hubungan korelasi antara dua variabel yang diukur dengan penskalaan ordinal. Dalam diskursus ini beberapa koefisien korelasi yang dapat dimanfaatkan adalah korelasi *Rank-Spearman*, *Gamma* dan *Tau-Kendal*. Tetapan ini merupakan

korelasi *nonparametrik*, yang tidak memerlukan distribusi dan lazimnya untuk melihat penjenjangan atau peringkat suatu variabel.

Rumusan koefisien korelasi *Rank-Spearman* yang terkoreksi adalah:

$$T_x = \frac{t_x^3 - t_x}{12}$$

$$T_y = \frac{t_y^3 - t_y}{12}$$

$$\sum X^2 = \frac{n^3 - n}{12} - \sum T_x$$

$$\sum y^2 = \frac{n^3 - n}{12} - \sum T_y$$

$$r_s = \frac{\sum x^2 + \sum y^2 - \sum d^2}{2\sqrt{\sum x^2 + \sum y^2}}$$

Di mana:

T_x adalah jumlah seri pada variabel x.

T_y adalah jumlah seri pada variabel y.

d adalah perbedaan antara rank variabel x dengan rank pada variabel y.

r_s adalah korelasi Spearman terkoreksi.

c. Data Kontinum

Jika peneliti bergerak dalam pengukuran nilai korelasi pada data kontinu, maka wajib dipenuhi beberapa asumsi yang di antaranya:

1. skala pengukuran berjenis interval;
2. memiliki sifat linearitas;
3. mengikuti distribusi normal;
4. memiliki sifat homoskedastisitas (varians homogen);
5. pengamatannya independen; dan
6. representatif sampling.

Tetapan atau koefisien korelasi yang dapat dilakukan untuk analisis untuk data kontinu adalah korelasi Pearson (*product moment*). Tetapan itu dapat diformulasikan sebagai berikut:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Statistik uji yang digunakan untuk menguji koefisien korelasi pearson adalah:

$$F = \frac{r^2/1}{(1-r^2)/n-2}$$

$$t = \sqrt{F}$$

$$t = r \sqrt{\frac{(n-2)}{(1-r^2)}}$$

$$df = n-2$$

Di mana:

r adalah koefisien korelasi pearson.

F adalah statistik uji F-Snadecor.

t adalah statistik uji t-student.

n adalah jumlah data.

df adalah derajat bebas (*degree of freedom*).

d. Data Deret Waktu (*Time Series*)

Dalam konteks data deret waktu, pada umumnya akan terjadi asosiasi atau ketergantungan antara nilai-nilai suatu deret berkala yang sama pada periode waktu yang berlainan yang disebut dengan *autokorelasi*. Sedangkan untuk menunjukkan hubungan antara nilai-nilai data deret waktu yang satu dengan deret waktu yang lainnya digunakan analisis *korelasi silang* (*cross correlation*). Statistik ini memiliki karakteristik yang sama seperti korelasi biasa (misalnya: *korelasi pearson*) yang nilainya berkisar antara -1 sampai +1. Di mana nilai koefisien korelasi +1 menunjukkan korelasi yang sempurna, nilai koefisien korelasi 0 menunjukkan tidak ada korelasi, dan nilai koefisien korelasi -1 menunjukkan anti-korelasi (korelasi sempurna/tinggi, tetapi kebalikan dari salah satu deret).

Misalkan untuk dua deret X_i dan Y_i di mana $i = 0, 1, 2, \dots, N-1$. Maka korelasi silang (r_{ks}) pada waktu keterlambatan sebesar d didefinisikan sebagai berikut:

$$r_{ks} = \frac{\sum_i [(X_i - \bar{X})x(Y_{i-d} - \bar{Y})]}{\sqrt{\sum_i (X_i - \bar{X})^2} \sqrt{\sum_i (Y_{i-d} - \bar{Y})^2}}$$

Di mana:

r_{ks} adalah besarnya koefisien korelasi silang antara X dengan Y

\bar{X} dan \bar{Y} adalah nilai rata-rata deret X_1 dan Y_1

d adalah waktu keterlambatan (*time lag*)

Jika persamaan (8.5) dihitung untuk semua $d=0, 1, 2, \dots, N-1$. Maka hasilnya adalah deret korelasi silang (r_d) dengan dua kali panjang deret-deret aslinya:

$$r_{ks} = \frac{\sum_i [(X_i - \bar{X})x(Y_{i-d} - \bar{Y})]}{\sqrt{\sum_i (X_i - \bar{X})^2} \sqrt{\sum_i (Y_{i-d} - \bar{Y})^2}}$$

Ilustrasi:

Apakah ada hubungan yang signifikan antara nilai statistik dengan nilai akuntansi dari para mahasiswa? Untuk itu diambil 12 orang mahasiswa dan dicatat nilainya seperti pada tabel di bawah.

Statistik	74	93	55	41	23	92	64	40	71	33	30	71
Akuntansi	81	86	67	35	30	100	55	52	76	24	48	97

Langkah untuk menjawab soal tersebut adalah: 1) Mencari nilai korelasi dan 2) melakukan pengujian apakah koefisien korelasi tersebut signifikan atau tidak?

Komputasi Koefisien Korelasi Linear Sederhana Antara Nilai Statistik dan Akuntansi dari 12 Mahasiswa

Statistik (X)	Akuntansi (Y)	XY	X ²	Y ²
74	81	5.994	5.476	6.561
93	86	7.998	8.649	7.396
55	67	3.685	3.025	4.489
41	35	1.435	1.681	1.225
23	30	690	529	900
92	100	9.200	8.464	10.000
64	55	3.520	4.096	3.025
40	52	2.080	1.600	2.704
71	76	5.396	5.041	5.776
33	24	792	1.089	576
30	48	1.440	900	2.304
71	87	6.117	5.041	7.569
Jumlah 687	741	48.407	45.591	52.525

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

$$r = \frac{12(48.407) - (687)(741)}{\sqrt{12(45.591) - (687)^2} \sqrt{12(52.525) - (741)^2}}$$

$$r = \frac{580.591 - 509.067}{\sqrt{12(45.591) - (687)^2} \sqrt{12(52.525) - (741)^2}}$$

$$r = \frac{71.817}{78.112,9} = 0,92$$

Proses uji hipotesis koefisien korelasi populasi dilakukan sebagai berikut:

1. Hipotesis $H_0 : \rho = 0$
 $H_1 : \rho \neq 0$
2. Uji statistik yang sesuai adalah uji t karena sampelnya kecil, yaitu 12 pasang.
3. Misalnya digunakan taraf signifikansi 0,05 dan 0,01
4. Nilai kritis tabel t adalah $\pm 2,228$ dan $\pm 3,169$
5. Harga uji statistik t dinyatakan dengan rumus: $t = r \sqrt{\frac{n-2}{1-r^2}}$

$$t = 0,92 \sqrt{\frac{12-2}{1-(0,92)^2}} = 7,42$$

6. Kesimpulan pengujian

Karena harga uji statistik $t = 7,42$ lebih besar daripada nilai kritis $t = 2,228$ dan $t = 3,169$, maka H_0 ditolak. Artinya, secara signifikan ada hubungan antara nilai statistik dengan nilai akuntansi, baik pada taraf signifikansi 5% maupun 1%.

Untuk sampel besar yakni apabila banyaknya pasangan 50 atau lebih, uji statistik yang sesuai adalah uji Z. Rumus untuk harga uji statistik Z adalah:

$$Z = \frac{r}{\frac{1}{\sqrt{n-1}}}$$

Sebagai ilustrasi, misalnya dari sampel yang terdiri dari 100 pasangan data diperoleh nilai $r = 0,45$. Apabila digunakan taraf signifikansi 0,05, pada pengujian dua sisi, maka nilai kritisnya adalah $\pm 1,96$.

Harga uji statistik:

$$Z = \frac{0,45}{\frac{1}{\sqrt{100-1}}} = 4,48$$

Dalam ilustrasi ini, H_0 ditolak pada taraf signifikansi 0,05 karena harga uji statistik $Z = 4,48$ jatuh pada daerah penolakan ($4,48 > 1,96$).

e. Analisis Korelasi Linear Parsial

Korelasi parsial digunakan untuk melihat hubungan tiga atau lebih variabel, di mana salah satu atau lebih variabel lainnya menjadi kontrol. Korelasi parsial dapat dilakukan pada data dengan skala minimal interval.

Pada analisis korelasi parsial digunakan untuk menyatakan hubungan antara variable X dengan variable Y, di mana variable Z sebagai kontrol (dianggap tetap). Nilai koefisien korelasi parsial antara X dengan Y jika Z dianggap tetap adalah:

$$r_{xy(Z)} = \frac{r_{xy} - r_{xy}r_{yy}}{\sqrt{(1-r_{xy}^2)(1-r_{yy}^2)}}$$

Untuk menguji koefisien korelasi parsial dari hipotesis berikut:

$H_0 : \rho_{YX} = 0$ vs $H_1 : \rho_{YX} \neq 0$, digunakan statistik uji-t, yaitu:

$$t = r_{YX_1(X_2)} \sqrt{\frac{n-3}{1-r_{YX_1(X_2)}^2}}$$

Dengan kriteria penolakan, tolak H_0 jika t -hitung $>$ t -tabel, dengan derajat bebas $n-3$, atau jika ρ -value $<$ α .

f. Analisis Korelasi Berganda (Multiple)

Analisis korelasi ganda digunakan untuk melihat keeratan hubungan antara variabel bebas yang ada X_1, X_2, \dots, X_k , secara bersama-sama terhadap variabel Y. Analisis korelasi berganda merupakan perluasan dari analisis korelasi sederhana. Dalam analisis regresi berganda bertujuan untuk mengetahui bagaimana derajat hubungan antara beberapa variabel independen (variabel X_1, X_2, X_k) dengan variabel dependen (Y) secara bersama-sama.

Nilai koefisien korelasi multipel antara Y dengan X_1, X_2, \dots, X_k , secara bersama-sama (simultan) adalah:

$$R_{y x_1 x_2 \dots x_k}^2 = \frac{b_{yx_1} \sum_{j=1}^n X_{1j} Y_j + b_{yx_2} \sum_{j=1}^n X_{2j} Y_j + \dots + b_{yx_k} \sum_{j=1}^n X_{kj} Y_j}{\sum_{j=1}^n Y_j^2}$$

Untuk menguji koefisien korelasi multipel dari hipotesis berikut:

$H_0 : \rho_{YX} = 0$ vs $H_1 : \rho_{YX} \neq 0$, digunakan statistik uji-F, yaitu:

$$F = \frac{R_{Y X_1 X_2 \dots X_k}^2 (n-k-1)}{(1-R_{Y X_1 X_2 \dots X_k}^2) 2}$$

Dengan kriteria penolakan, tolak H_0 jika F-hitung $> F$ -tabel, dengan derajat pembilang k dan derajat bebas penyebut n-k-1, atau jika ρ -value $< \alpha$.

Berdasarkan contoh tentang hubungan antara penghasilan keluarga X_1 dan besar keluarga X_2 dengan pengeluaran untuk bahan makanan Y, koefisien korelasi linear bergandanya dinyatakan dengan rumusan berikut.

$$R_{Y.12\dots k} = \sqrt{\frac{b_1 \sum X_1 Y + b_2 \sum X_2 Y}{\sum Y^2}}$$

Di mana:

$$b_1 = 0,0544; b_2 = 0,2304; \sum X_1 Y = 16,858; \sum X_2 Y = 9,96; \sum Y^2 = 3,777$$

Jadi koefisien korelasi berganda dari contoh tersebut adalah:

$$R_{Y.12\dots k} = \sqrt{\frac{0,0544(16,858) + 0,2304(9,96)}{3,777}} = \sqrt{\frac{3,2118592}{3,777}} = 0,92$$

sedangkan koefisien determinasi berganda (R^2) dari contoh tersebut adalah:

$$R_{Y.12\dots k} = \frac{b_1 \sum X_1 Y + b_2 \sum X_2 Y}{\sum Y^2} \text{ atau } \frac{SSR}{SST}$$

$$R_{Y.12\dots k} = \frac{0,0544(16,858) + 0,2304(9,96)}{3,777}$$

$$R_{Y.12\dots k} = \frac{3,2118592}{3,777} = 0,85 \text{ atau } (0,92)^2$$

Angka tersebut menunjukkan bahwa sekitar 85% dari variasi pengeluaran untuk bahan makanan Y dijelaskan oleh kombinasi dari penghasilan keluarga X_1 dan besar keluarga X_2 . Sisanya yakni 15% dijelaskan oleh variabel independen lainnya yang tidak teramati.

Dari kasus korelasi linear berganda, peneliti dapat menghitung koefisien korelasi parsialnya. Korelasi parsial (*partial correlation*) adalah korelasi antara sebuah variabel dependen Y dengan sebuah variabel independen X_1 , sementara sejumlah variabel independen lainnya konstan.

Apabila variabel independennya ada dua buah, yaitu X_1 dan X_2 maka koefisien korelasi parsial yang ada ialah $r_{y1.2}$ dan $r_{y2.1}$, yang masing-masing menunjukkan koefisien korelasi antara Y dengan X_1 apabila X_2 konstan dan koefisien korelasi antara Y dengan X_2 apabila X_1 konstan. Seperti dalam contoh tersebut di muka, $r_{y2.1}$, menunjukkan korelasi antara besar keluarga X_2 dengan pengeluaran untuk bahan makanan Y apabila penghasilan keluarga X_1 konstan.

Rumus-rumusnya adalah:

$$R_{y1.2} = \frac{r_{y1} - (r_{y2})(r_{12})}{(1 - r_{y2}^2)(1 - r_{12}^2)}$$

$$R_{y2.1} = \frac{r_{y2} - (r_{y1})(r_{12})}{(1 - r_{y1}^2)(1 - r_{12}^2)}$$

Untuk menghitung koefisien korelasi parsialnya terlebih dahulu harus dihitung koefisien korelasi sederhana antara X_1 dengan Y, antara X_2 dengan Y, dan antara

X_1 dengan X_2

Berdasarkan contoh dimuka :

$$R_{x1y} = \frac{\sum X_1y}{\sqrt{(\sum X_1^2)(\sum Y^2)}} = \frac{16,858}{\sqrt{(438,416)(3,77)}} = \frac{16,858}{40,693} = 0,41$$

$$R_{x2y} = \frac{\sum X_2y}{\sqrt{(\sum X_2^2)(\sum Y^2)}} = \frac{9,96}{\sqrt{(50,4)(3,777)}} = \frac{9,96}{13,797} = 0,72$$

$$R_{x1x2} = \frac{\sum X_1X_2}{\sqrt{(\sum X_1^2)(\sum X_2^2)}} = \frac{-30,38}{\sqrt{(438,416)(50,4)}} = \frac{9,96}{148,65} = -0,20$$

Koefisien korelasi parsial :

$$R_{y1.2} = \frac{0,41 - (0,72)(-0,20)}{(1 - 0,72^2)(1 - 0,20^2)} = \frac{0,554}{(0,694)(0,98)} = 0,81$$

$$R_{y2.1} = \frac{0,72 - (0,41)(-0,20)}{(1 - 0,41^2)(1 - 0,20^2)} = \frac{0,802}{(0,912)(0,98)} = 0,90$$

SOAL – SOAL LATIHAN

1. Apa yang dimaksud dengan analisis korelasi dan tetapan/koefisien korelasi dalam penelitian kuantitatif?
2. Tetapan korelasi berada dalam ambang 0,00 hingga 1,00 yang terbagi dalam lima interval untuk menjelaskan kekuatan hubungan. Berikan penjelasan mengenai penafsiran atas koefisien korelasi tersebut?
3. Bagaimana rumusan tetapan koefien korelasi untuk data yang diambil dengan skala ordinal?
4. Apakah asumsi – asumsi yang harus dipenuhi peneliti jika ingin mencari korelasi pada data yang di himpun dengan penskalaan kontinu?
5. Dalam pengukuran terhadap 12 mahasiswa fakultas ekonomi universitas X, pada mata kuliah matematika diskrit dan ekonometri, diperoleh data sebagai berikut :

Matematika Diskrit	76	92	45	55	33	82	74	40	61	33	48	76
Ekonometri	80	86	57	45	37	90	75	42	66	44	43	77

Pertanyaannya, apakah ada hubungan yang signifikan antara nilai matematika diskrit dengan nilai ekonometri pada mahasiswa tersebut?

BAB VIII

ANALISIS REGRESI LINEAR

8.1 Analisis Regresi Linear Sederhana

Jika kita perhatikan, tujuan utama dari kebanyakan penelitian dalam bidang ekonomi dan bisnis adalah mengadakan pendugaan atau peramalan. Misalnya, ramalan tentang besarnya tingkat penjualan berdasarkan besarnya biaya promosi, ramalan tentang permintaan pasar suatu komoditi berdasarkan tingkat harga komoditi tersebut. Peramalan tentang perubahan harga sekuritas berdasarkan perubahan dalam beberapa indeks pasar dan sebagainya.

Hubungan antara variabel-variabel, yaitu antara variabel yang telah diketahui dengan variabel yang akan diramalkan diformulasikan dalam bentuk persamaan matematis. Analisis regresi adalah merupakan prosedur di mana dengan melalui formulasi persamaan matematis, hendak diramalkan nilai variabel random kontinu berdasarkan nilai variabel kumulatif lainnya yang diketahui. Variabel yang nilainya hendak diduga berdasarkan persamaan regresi disebut variabel dependen, dan variabel yang digunakan sebagai dasar untuk membuat pendugaan disebut variabel independen.

Hubungan antara independen variabel dengan dependen variabel tersebut merupakan hubungan linear dan dinyatakan dengan model persamaan matematis linear. Jika dinyatakan dengan model persamaan matematis linear. Jika nilai variabel dependen diduga berdasarkan satu variabel independen saja, analisisnya dinamakan analisis regresi sederhana (*Simple Regression Analysis*). Adapun apabila nilai variabel dependen diduga berdasarkan dua atau lebih variabel independen, analisisnya dinamakan analisis regresi berganda (*Multiple Regression Analysis*).

Dalam analisis regresi tersebut, hubungan antara variabel independen dengan Variabel dependen merupakan hubungan linear, dan hubungan ini merupakan hubungan statistikal, artinya tidak ada variabel dependen yang pasti untuk setiap variabel independen yang diketahui. Apabila nilai variabel dependen dapat diramalkan dengan pasti berdasarkan nilai variabel independen yang diketahui, dinamakan mempunyai hubungan fungsional. Studi hubungan variable-variabel ekonomi atau bisnis pada umumnya merupakan hubungan statistik.

Analisis regresi linear sederhana tersebut didasarkan pada asumsi-asumsi sebagai berikut:

1. Variabel independen dan variabel dependen mempunyai hubungan linear. Persamaan linearnya dinyatakan dengan: $Y_i = \alpha + \beta x + e_i$

Di mana:

Y = Nilai variabel dependen dalam observasi atau percobaan ke-1.

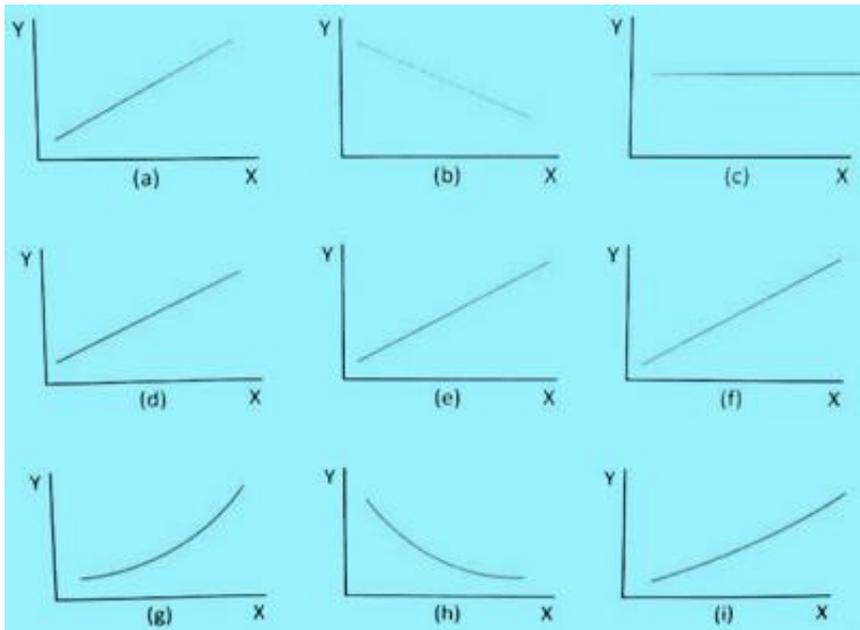
A = Parameter pertama dari persamaan regresi yang menunjukkan nilai Y apabila

X = 0.

B = Parameter kedua dari persamaan regresi menunjukkan slope dari garis regresi.

e_i = Random *error* dalam observasi atau percobaan ke-1, berkaitan dengan proses sampling.

2. Variabel dependen merupakan variabel random kontinu. Sedangkan variabel independennya merupakan serangkaian nilai yang ditentukan atau diketahui dan bukan random. Variabel X merupakan variabel yang nilainya sudah dirancang, sedangkan variabel dependen Y nilainya diperoleh melalui proses sampling. Dan nilai e_i merupakan sampling error yang berhubungan dengan variabel random dependen.
3. Distribusi kondisional variabel dependen, untuk berbagai nilai variabel independen tertentu semua berdistribusi normal. Dalam model regresi linear, asumsi ini menunjukkan bahwa distribusi sampling error e_i juga normal.
4. Variance dari distribusi kondisional variabel dependen, untuk berbagai nilai variabel independen tertentu, semuanya sama atau statis dan ini berarti bahwa variance yang berkaitan dengan sampling error e_i adalah sama untuk berbagai nilai X_i asumsi ini dinamakan *homoscedasticity*.
5. Nilai observasi yang satu dengan yang lain dari variabel random, tidak berkorelasi (*uncorrelated*). Asumsi ini juga menunjukkan bahwa jumlah-jumlah random *error* e_i juga tidak berkorelasi. Untuk mengetahui apakah pasangan nilai variabel independen dan variabel dependen mempunyai hubungan linear atau tidak, perlu dibuat diagram pencar (scatter diagram) seperti dibawah ini.



Gambar: Diagram berserak untuk berbagai tipe hubungan: a) hubungan linear positif; b) hubungan linear negative; c) tidak berhubungan; d) hubungan linear positif dengan derajat rendah; e) hubungan linear positif dengan derajat tinggi; f) hubungan linear positif dengan mengabaikan asumsi *homoscedasticity*; g) hubungan curvilinear positif; h) hubungan curvilinear negative; dan i) hubungan positif dengan mengabaikan asumsi *homoscedasticity*.

Parameter α dan β dalam model regresi linear seperti pada asumsi pertama, diduga dengan nilai a dan b yang dihitung dari data sampel. Persamaan regresi linear untuk menduga nilai variabel dependen (Y') berdasarkan nilai variabel independen (X) tertentu, dinyatakan dengan:

$$Y' = a + bX$$

Nilai b (slope garis regresi) dicari dengan rumus

$$b = \frac{\sum xy}{\sum x^2}$$

dimana :

$$\sum x^2 = \sum_{i=1}^n (X_i - \bar{X})^2$$

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

$$\sum XY = \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})^2$$

Nilai a (intersep garis regresi) dicari dengan rumus:

$$a = \frac{\sum_{i=1}^n Y_i - b(\sum_{i=1}^n X_i)}{n} \quad \text{atau} \quad a = \bar{Y} - b(\bar{X})$$

Ilustrasi:

Para peneliti bidang produksi pertanian bermaksud untuk mengetahui bagaimana pengaruh dari berbagai tingkat penggunaan pupuk NPK Phonska terhadap hasil panen jagung di suatu kabupaten X. Dari sampel empat tingkatan penggunaan pupuk, diperoleh data sebagai berikut:

Tingkat Penggunaan Pupuk NPK Phonska Kg/ Ha	Hasil Panen Jagung Kg / Ha
0	4.230
50	5.442
100	6.661
150	7.150

Persamaan regresi linear, $Y' = a + bX$, dengan metode kuadrat terkecil (*least-squares*) dihitung dengan bantuan tabel seperti berikut ini.

Tingkat Penggunaan Pupuk NPK Phonska Kg/ Ha	Hasil Panen Jagung Kg / Ha	XY	X ²	Y ²
0	4.230	0	0	17.892.900
50	5.442	272.100	2.500	29615.364
100	6.661	666.100	10.500	44.368.921
150	7.150	1.072.500	22.500	51.122.500
Jumlah 300	23.483	2.010.700	35.000	142.999.685
Mean 75	5.870,75			

Tingkat Penggunaan Pupuk Phonska Kg/ Ha	Hasil Panen Jagung Kg/ Ha	Deviasi dari Mean	Deviasi Kuadrat	(X)(Y)		
0	4.230	-75	1.660,75	5.625	2.692.061	123.056
50	5.442	-25	-425,75	625	183.827	10.719
100	6.661	25	790,25	625	624.495	19.756
150	7.150	75	1.279,25	5.625	1.636.481	95.944
Jumlah 300	23.483	0	0,00	12.500	5.136.864	249.475
Mean 75	5.870,75					

Slope garis regresi (b) dihitung sebagai berikut:

Rumus (1):

$$b = \frac{4(2.010.700) - (300)(23.483)}{4(35.000) - (300)^2} = \frac{8.042.800 - 7.044.900}{140.000 - 90.000} \quad b = \frac{997.900}{50.000} = 19,96$$

Rumus (2):

$$b = \frac{4(2.010.700) - 4(75)(5.870,75)}{(35.000) - 4(75)^2} = \frac{2.010.700 - 1.761.225}{35.000 - 22.500} \quad b = \frac{249.475}{12.500} = 19,96$$

Rumus (3)

$$b = \frac{249.475}{12.500} = 19,96$$

Intersep garis regresi (a) dihitung sebagai berikut;

Rumus (1)

$$a = \frac{23.483 - 19,96(300)}{4} = \frac{23.483 - 5.989}{4} = 4.374$$

Rumus (2)

$$a = 5.870,75 - 19,96(75) = 5.870,75 - 1.497 = 4.374$$

Jadi, persamaan regresi linearnya adalah:

$$Y' = 4.374 + 19,96 X \text{ untuk } (0 \leq X \leq 150)$$

Intersep (a) sama dengan 4.374 artinya apabila X=0 (Tidak dilakukan pemupukan dengan NPK Phonska), hasil panen jagung sebesar 4.374 kh/ha.

Slope garis regresi (b) sama dengan 19,96 artinya bahwa setiap penambahan pupuk NPK Phonska 1 kg/ha akan memberi tambahan hasil panen padi sebanyak 19,96 kg/ha.

Persamaan linear yang diperoleh bisa digunakan untuk memprediksi panen jagung yang diharapkan. Misalnya apabila tingkat pemupukan nitrogen 0 kg/ha, 50 kg/ha, 100 kg/ha, 150 kg/ha, berapa hasil panen jagung yang diharapkan?

$$Y' = 4.374 + 19,96 (0) = 4.374 \text{ kg/ha.}$$

$$Y' = 4.374 + 19,96 (50) = 5.372 \text{ kg/ha.}$$

$$Y' = 4.374 + 19,96 (100) = 6.370 \text{ kg/ha.}$$

$$Y' = 4.374 + 19,96 (150) = 7.368 \text{ kg/ha.}$$

Besarnya angka peramalan atau pendugaan tersebut adalah rata-rata (*mean*) dari berbagai kemungkinan jumlah hasil panen, berdasarkan tingkat pemupukan tertentu. Dalam kegiatan analisis regresi linear, yang hitung adalah β dan α . Hipotesis nihil yang diuji mengatakan bahwa $\beta = 0$, artinya tidak terdapat hubungan (*relationship*) antara variabel independen dengan variabel dependen; dan bahwa $\alpha = \alpha_0$ artinya intersep regresi linear ($X = 0$) sama dengan harga tertentu.

Seperti uji hipotesis yang telah dibicarakan sebelumnya, proses pengujian β dan α tersebut dilakukan melalui enam tahap.

Sehubungan dengan contoh tersebut, dengan pengujian β dilakukan sebagai berikut:

1. Rumusan Hipotesis

$H_0 : \beta = 0$; artinya tidak ada pengaruh yang signifikan antara variabel bebas terhadap variabel respon/variabel dependen/variabel terikat.

$H_1 : \beta \neq 0$; artinya ada pengaruh yang signifikan antara variabel bebas terhadap variabel respon/variabel dependen/variabel terikat.

2. Uji statistik yang digunakan adalah uji t

3. Misalnya digunakan taraf signifikansi 0,05. Untuk ukuran sampel, dalam contoh tersebut terdiri dari 4 pasang data ($n - 4$).

4. Nilai kritis untuk menentukan daerah penolakan, yakni menurut tabel nilai $t = \pm 4,303$.

5. Perhitungan harga uji statistis t adalah sebagai berikut:

$$t_b = \frac{b - \beta_0}{S_b}$$

Karena dihipotesiskan bahwa $\beta = 0$ maka rumusnya menjadi:

$$t_b = \frac{b}{S_b}$$

S_b yang merupakan penyimpangan baku dari koefisien regresi, dihitung dengan rumus:

$$S_b = \frac{S_{yx}}{\sum_{i=1}^n X_1^2 - \frac{(\sum_{i=1}^n X_1)^2}{n}} \quad \text{atau} \quad S_b = \frac{S_{yx}}{\sqrt{\sum X^2}}$$

Sedang S_{yx} merupakan penyimpangan baku dari pendugaan nilai Y, yaitu deviasi standar kondisional dari variabel dependen (Y) berdasarkan nilai variabel independen (X) tertentu, dinyatakan dengan rumus:

$$S_{yx} = \sqrt{\frac{\sum(Y-Y')^2}{n-2}} \quad \text{atau} \quad S_{yx} = \sqrt{\frac{\sum y^2 - \frac{(\sum xy)^2}{\sum x^2}}{n-2}}$$

Untuk menghitung “*standard error of estimate*” berdasarkan rumus ke-1, digunakan tabel bantu sebagai berikut.

Y	Y'	(Y-Y')	(Y-Y') ²
4.230	4.374	-144	20.736
5.442	5.372	70	4.900
6.661	6.370	291	84.681
7.150	7.368	-218	47.524
			157.524

$$S_{yx} = \sqrt{\frac{\sum(Y-Y')^2}{n-2}}$$

$$S_{yx} = \sqrt{\frac{157.841}{2}} = \sqrt{78.921} = 280,93$$

Dari contoh di atas, berdasarkan rumus “*standard error of estimate*” yang ke-2 adalah:

$$S_b = \frac{S_{yx}}{\sqrt{\sum X^2}} = \frac{280,93}{\sqrt{12.500}} = \frac{280,93}{111,80} = 2,51$$

Maka harga uji t:

$$t_b = \frac{19,96}{2,51} = 7,95$$

6. Kesimpulan pengujian

Nilai $t = 7,95$ dalam uji, ternyata lebih besar daripada nilai kritis $t = 4,303$, yang artinya berada di daerah penolakan, maka H_0 ditolak. Maknanya memang ada hubungan positif antara tingkat penggunaan pupuk NPK Phonska dengan hasil panen jagung pada taraf signifikansi 0,05.

Untuk pengujian α , dimisalkan bahwa hasil panen jagung bila tidak dilakukan pemupukan adalah 4.000kg/ha.

Hipotesis:

$$H_0 : \alpha = 4.000$$

$$H_1 : \alpha \neq 4.000$$

Harga uji 96statistic t, digunakan rumus:

$$t = \frac{a - a_0}{S_{yx} \sqrt{\left[\frac{1}{n} + \frac{X^2}{\sum X^2} \right]}}$$
$$t = \frac{4.374 - 4.000}{280,93 \sqrt{\left[\frac{1}{4} + \frac{(75)^2}{12.500} \right]}} = \frac{374}{235} = 1,59$$

Seperti di ketahui nilai uji statistik $t = 1,59$ lebih kecil daripada nilai kritis $t = 4,303$ (berada di daerah penerimaan), maka H_0 diterima pada taraf signifikansi 0,05. Intersep $\alpha = 4.374$ kg/ha tidak berbeda secara signifikan dengan 4.000 kg/ha.

8.2 Analisis Regresi Linear Berganda

Secara teoretis, analisis regresi linear berganda merupakan pengembangan dari analisis regresi linear sederhana. Dalam konteks analisis regresi linear sederhana dibuat analisis hubungan satu variabel independen dengan satu variabel dependen. Hubungan itu dinyatakan dalam persamaan linear yang dimaksudkan untuk meramalkan besarnya nilai Y berdasarkan nilai X tertentu.

Dalam praktik, kemampuan untuk memprediksi perubahan variabel dependen (Y) akan menjadi lebih baik jika dimasukkan lebih dari satu variabel independen dalam persamaan linearnya. Korelasi antara lebih dari satu variabel bebas dengan satu variabel terikat inilah yang dibahas dalam analisis regresi linear berganda. Hal ini sesuai dengan realitas bahwa relasi antar variabel dalam kajian ilmu sosial bersifat statistikal, yang maksudnya

adalah perubahan nilai Y tidak secara mutlak hanya dipengaruhi oleh satu nilai X tertentu dan tunggal, namun akan dipengaruhi oleh beberapa besaran nilai X. Sebagaimana pada analisis regresi linear sederhana, dalam analisis regresi linear berganda asumsi-asumsinya adalah:

1. Variabel independen dan dependen berhubungan secara linear, yang dinyatakan dengan model:

$$Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + e_i$$

Di mana:

Y_i	Nilai variabel dependen hasil observasi pertama
α	Nilai Y bila semua X = 0
β_j	Koefisien regresi berhubungan dengan masing-masing variabel independen X
X_{ij}	Nilai variabel independen ke-j dari observasi ke-i, di mana j = k
e_i	Random error dari observasi ke-i

2. Dalam analisis ini, variabel dependennya adalah variabel acak kontinu, sedangkan variabel-variabel independennya bersifat sudah tertentu, telah diketahui dan non random. Nilai e, dalam model linear merupakan sampling error bagi variabel acak dependen.
3. Dalam analisis regresi linier berganda, distribusi kondisional variabel dependen, semuanya berdistribusi normal. Hal ini berarti distribusi sampling error e dalam model ini juga normal.
4. Asumsi berikutnya adalah varian dari distribusi kondisional variabel dependen, semuanya sama atau bersifat homogeni.
5. Harga atau nilai observasi dari variabel acak yang satu dengan yang lainnya tidak saling berkaitan, sehingga besarnya random error e dalam model linear juga tidak berhubungan.

Rumusan matematik analisis regresi linear berganda, berdasarkan penelitian sampel dinyatakan sebagai berikut:

$$Y_i = a + B_1 X_{i1} + B_2 X_{i2} + \dots + B_k X_{ik}$$

Dari rumusan diatas, Y merupakan variabel yang akan diprediksi dan X merupakan variabel yang telah diketahui untuk membantu memprediksi. Untuk memberikan gambaran bagaimana membuat analisis regresi linear berganda, berikut adalah ilustrasinya:

Persoalan mendasar dalam hidup salah satunya adalah ketersediaan makanan. Seorang peneliti berpendudukan ingin memetakan kebutuhan

keluarga dalam pengeluaran untuk pembelanjaan bahan makanan. Peneliti itu akan memprediksikan besarnya pengeluaran untuk bahan makanan per bulan (Y) Berdasarkan besarnya penghasilan keluarga per bulan (X_1) dan banyaknya keluarga (X_2). Berdasarkan sampel acak 15 keluarga diperoleh informasi sebagai berikut:

Penghasilan Keluarga (Ratusan Ribu Rp/Bulan) X_1	Besar Keluarga X_2	Pengeluaran Bahan Makanan (Ratusan Ribu Rp/Bulan) Y
5,5	1	0,8
8,9	1	1,0
21,9	1	1,7
6,8	2	1,4
7,5	2	1,2
1,72	2	1,8
2,21	2	1,9
1,90	3	2,3
1,20	3	1,7
1,40	4	1,5
1,09	4	1,8
7,5	5	2,0
1,40	5	2,2
1,37	6	2,8
6,0	7	2,2

Untuk kasus dua variabel independen, persamaan linearnya dinyatakan dengan:

$$Y_t = a + b_1X_t + b_2X_{2t}$$

Untuk mendapatkan nilai a, b_1 dan b_2 digunakan rumus-rumus sebagai berikut:

$$b_1 = \frac{(\sum X_2^2)(\sum x_2y) - (\sum X_1X_2)(\sum X_2y)}{(\sum X_1^2)(\sum x_2^2) - (X_1X_2)^2}$$

$$b_2 = \frac{(\sum X_1^2)(\sum x_2y) - (\sum X_1X_2)(\sum X_1y)}{(\sum X_1^2)(\sum x_2^2) - (X_1X_2)^2}$$

$$a = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2$$

Di mana:

$$\sum X_1^2 = \sum X_1^2 - \frac{(\sum X_1)^2}{n}$$

$$\sum X_2^2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n}$$

$$\sum X^2 Y = \sum X^2 Y - \frac{(\sum X_1)(\sum Y)}{n}$$

$$\bar{Y} = \frac{\sum Y}{n}$$

$$\bar{X}_1 = \frac{\sum X_1}{n}$$

$$\bar{X}_2 = \frac{\sum X_2}{n}$$

Tabel: Komputasi Persamaan Regresi Linear Berganda Berdasarkan Hubungan Variabel Penghasilan Keluarga (X₁) dan Variabel Besar Keluarga (X₂) dengan Variabel untuk Bahan Makanan (Y)

X ₁	X ₂	Y	X ₁ ²	X ₂ ²	Y ²	X ₁ Y	X ₂ Y	X ₁ X ₂
5,5	1	0,8	30,25	1	0,64	4,40	0,8	5,5
8,9	1	1,0	79,21	1	1,00	8,90	1,0	8,9
21,9	1	1,7	475,24	1	2,89	31,06	1,7	21,8
6,8	2	1,4	46,24	4	1,96	9,52	2,8	13,6
7,5	2	1,2	56,25	4	1,44	9,00	2,4	15,0
1,72	2	1,8	295,84	4	3,24	30,96	3,6	34,4
2,21	2	1,9	488,41	4	3,61	41,99	3,8	44,2
1,90	3	2,3	361,00	9	5,29	43,70	6,9	57,0
1,20	3	1,7	144,00	9	2,89	20,40	5,1	36,0
1,40	4	1,5	196,00	16	2,25	21,00	6,0	56,0
1,09	4	1,8	118,81	16	3,24	19,62	7,2	43,6
7,5	5	2,0	56,25	25	4,00	15,00	10,0	37,5
1,40	5	2,2	196,00	25	4,84	30,80	11,0	70,0
1,37	6	2,8	187,69	36	7,84	38,36	16,8	82,2
6,0	7	2,1	36,00	49	4,41	12,60	14,7	42,0
186,9	48	26,2	2.767,19	204	49,54	343,31	93,8	567,7
Mean 12,46	3,2	1,74						

$$b_1 = \frac{(50,4)(16,858) - 30,38}{(9,96)(438,416) - (-30,38)^2}$$

$$b_1 = \frac{849,6432 - 302,5848}{22.096,1664 - 922,944}$$

$$b_1 = \frac{1.152,228}{21.173,222} = 0,05442$$

$$b_2 = \frac{(438,416)(9,96) - 30,38}{(16,858)(438,416)(50,4) - (-30,38)^2}$$

$$b_2 = \frac{4.366,62336 - 512,14604}{22.096,1664 - 922,9444}$$

$$b_2 = \frac{8.828,7694}{21.173,222} = 0,23042$$

Intersepnya adalah:

$$a = 1,74 - 0,05442(12,46) - 0,23042(3,2)$$

$$a = 1,74 - 0,6780732 - 0,737344$$

$$a = 1,74 - 1,4154172 = 0,3245828$$

Persamaan regresi linear bergandanya adalah:

$$Y' = 0,3246 + 0,0544X_1 + 0,2304X_2$$

Pengertian dari persamaan tersebut adalah: pertama apabila X_2 konstan, penambahan satu unit pada X_1 , akan mempunyai pengaruh menaikkan 0,0544 unit pada Y . Kedua apabila X_1 konstan, penambahan satu unit pada X_2 akan mempunyai pengaruh menaikkan 0,2304 unit pada Y . Ketiga apabila X_1 dan X_2 sama dengan nol, besarnya Y adalah 0,3246 satuan.

Berdasarkan persamaan tersebut dapat dibuat prediksi nilai-nilai Y berdasarkan kombinasi nilai X_1 dan X_2 tertentu.

$$Y_1' = 0,3246 + 0,0544(5,5) + 0,2304(1) = 0,8542$$

$$Y_2' = 0,3246 + 0,0544(8,9) + 0,2304(1) = 1,0392$$

...

$$Y_{15} = 0,3246 + 0,0544(6,0) + 0,2304(7) = 2,2638$$

Dalam analisis regresi linear berganda dalam contoh tersebut, hipotesis nihil yang akan diuji mengatakan bahwa variabel penghasilan keluarga X_1 dan variabel besar keluarga X_2 secara bersama-sama tidak mempengaruhi perubahan variabel pengeluaran untuk bahan makanan Y . Secara umum dikatakan, bahwa variabel-variabel independen yang diamati secara bersama-sama tidak mempunyai pengaruh yang signifikan terhadap variabel dependen.

Dalam analisis regresi parsial (*partial regression coefficient*), uji signifikansi dibuat untuk masing-masing koefisien regresi. Misalnya dalam contoh tersebut, hipotesis nihilnya mengatakan bahwa $\beta = 0$ dengan catatan

X_2 dianggap konstan, dan β_2 dengan catatan bahwa X_1 dianggap konstan, secara umum dinyatakan bahwa:

$$H_0 : \beta_i = 0$$

Untuk kasus regresi linear berganda, dapat pula dibuat analisis regresi sederhana (*simple regression* dari Y atas X_1 dan Y atas X_2) untuk membandingkan antara X_1 dengan X_2 mana yang merupakan yang lebih baik dari Y . Pengujian hipotesis didahului menghitung terlebih dahulu *standard error of estimate* dengan menggunakan alat bantu seperti berikut ini :

Tabel. Komputasi $(Y-Y')^2$ dan *Standard Error of Estimate*

X_1	X_2	Y	Y'	$(Y-Y')$	$(Y-Y')^2$
5,5	1	0,8	0,8542	-0,0542	0,002938
8,9	1	1,0	0,0392	-0,0392	0,01537
21,9	1	1,7	1,7409	-0,0409	0,001673
6,8	2	1,4	1,1553	0,2447	0,059878
7,5	2	1,2	1,1934	0,0066	0,000044
1,72	2	1,8	1,7211	0,0078	0,006225
2,21	2	1,9	1,9876	0,0876	0,007674
1,90	3	2,3	2,0494	0,2506	0,062800
1,20	3	1,7	1,6686	0,0314	0,000986
1,40	4	1,5	2,0078	-0,5078	0,257861
1,09	4	1,8	1,8392	-0,0392	0,001537
7,5	5	2,0	1,8864	0,1154	0,013317
1,40	5	2,2	2,2382	-0,0382	0,001459
1,37	6	2,8	2,4523	-0,3477	0,120895
6,0	7	2,1	2,2638	-0,1638	0,026830
186,9	48	26,2			0,565654

Standard error of estimate dinyatakan dengan rumus:

$$S_{y.12} = \sqrt{\frac{\sum (Y-Y')^2}{n-3}} \text{ atau } S_{y.12} = \sqrt{\frac{\sum y^2 - b_1 \sum X_1 Y - b_2 \sum X_2 Y}{n-3}}$$

Berdasarkan contoh tersebut di muka, besarnya *standard error of estimate* adalah:

Rumus (1):

$$S_{y.12} = \sqrt{\frac{0,565654}{15-3}} = \sqrt{0,0471378} = 0,217$$

Rumus (2):

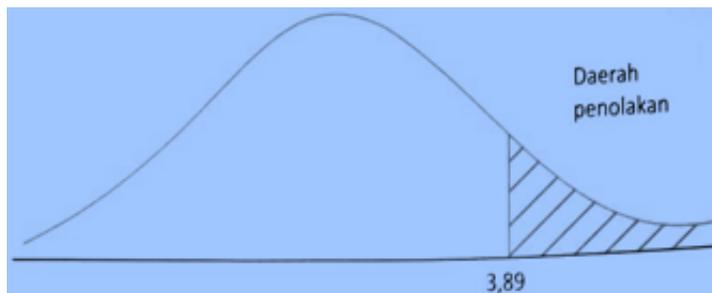
$$S_{y.12} = \sqrt{\frac{3,777 - 0,0544(16,858) - 0,2304(9,96)}{15-3}}$$

$$S_{y.12} = \sqrt{\frac{3,777 - 0,9170752 - 2,294784}{12}}$$

$$S_{y.12} = \sqrt{\frac{0,56514}{12}} = \sqrt{0,047095} = 0,217$$

Proses pengujian hipotesis koefisien regresinya dilakukan sebagai berikut:

1. Formulasi hipotesis nihil dan hipotesis :
 H_0 : penghasilan keluarga X_1 dan besar keluarga X_2 secara bersama-sama tidak berpengaruh secara signifikan terhadap besarnya pengeluaran untuk bahan makanan Y .
 H_1 : penghasilan keluarga X_2 dan besar keluarga X_2 secara bersama-sama mempunyai pengaruh yang signifikan terhadap besarnya pengeluaran untuk bahan makanan Y .
2. Uji statistik yang sesuai adalah uji F.
3. Misalnya taraf signifikansi yang digunakan adalah 0,05. Dan dalam contoh tersebut besar sampel $n = 15$
4. Nilai kritis $F = 3,89$.



Gambar 9.2 Daerah Penolakan F

5. Harga uji dihitung dengan rumus:

$$F = \frac{\frac{SSR}{df}}{\frac{SSE}{df}} = \frac{\frac{SSR}{k}}{SSE(n-k-1)}$$

Berdasarkan contoh di muka:

$$\begin{aligned} SSR &= 0,544(16,858) + 0,2304(9,96) \\ &= 0,9170752 + 2,2947840 \end{aligned}$$

$$= 3,2118592$$

$$SSE = 0,565654$$

$$SST = \sum Y^2 = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

$$= 49,54 - \frac{(26,2)^2}{2} = 3,777$$

$$SSE = 3,777 - 3,2118592 = 0,5651408$$

(ada perbedaan karena adanya pembulatan)

$$F = \frac{\frac{3,2118592}{2}}{\frac{0,565654}{(15-2-1)}}$$

$$F = \frac{1,6059296}{0,04731378} = 34,07$$

6. Kesimpulan, oleh karena harga uji $F = 34,07$ lebih besar dari harga kritis $F = 3,89$, maka H_0 ditolak (H_1 diterima) pada taraf signifikansi 0,05. Dengan demikian penghasilan keluarga dan besar keluarga secara bersama berpengaruh signifikan terhadap pengeluaran untuk makanan.

Selanjutnya, proses pengujian partial *regression coefficient* dilakukan sebagai berikut:

Untuk β_1 :

1. Hipotesis: $H_0 : \beta_1 = 0$

$$H_1 : \beta_1 \neq 0$$

2. Uji sttatistik yang sesuai adalah uji t.
3. Misalnya digunakan taraf signifikansi 0,05. Dalam contoh $n - 15$.
4. Nilai kritis $t = 2,179$.
5. Harga uji t:

$$t = \frac{b^1 - \beta_1}{sb_1}$$

$$S_{b_1} = \frac{S_{y_{12}}}{\sqrt{\sum x_1^2 (1 - r_{12}^2)}}$$

Dimana:

$S_{y_{12}}$ = Standart error of estimasi

R_{12} = koefisien korelasi sederhana antara X_1 dan X_2

$$r_{12} = \frac{-30,38}{\sqrt{(438,416)(50,4)}} = -0,20$$

$$S_{b1} = \frac{0,217}{\sqrt{2767,19[1-(0,20)]^2}}$$

$$= \frac{0,217}{\sqrt{2656,5024}} = \frac{0,217}{51,64} = 0,0042$$

$$t = \frac{0,0544 - 0}{0,0042} = 12,952$$

6. Kesimpulan pengujian:

Harga uji $t = 12,952$ lebih besar dibanding nilai kritis $t = 2,179$ maka H_0 ditolak pada derajat signifikansi 0,05. Dengan menganggap besar keluarga X_2 konstan. Penghasilan keluarga X_1 mempunyai pengaruh yang signifikan terhadap pengeluaran untuk bahan makanan Y . Apabila ingin dibuat analisis regresi sederhana, uji hipotesisnya dilakukan sebagai berikut: Regresi sederhana dari Y atas X_1 :

$$b = \frac{\sum X_1 Y}{\sum X_1^2} = \frac{16,858}{438,416} = 0,03845 \quad a = \bar{Y} - b\bar{X}_1 = 1,74 - 0,03845(12,46) = 1,2609$$

$$Y' = 1,2609 + 0,03845X_1$$

$$SSR = \frac{(\sum X_1 Y)^2}{\sum X_1^2} = \frac{(16,858)^2}{438,416} = 0,648225$$

$$SSE = 3,777 - 0,648225 = 3,128775$$

$$R_{y,1}^2 = \frac{SSR}{SST} = \frac{0,648225}{3,777} = 0,1716 \quad (\text{koefisien determinasi})$$

$$R_y = \sqrt{R_{y,1}^2} = \sqrt{0,1716} = 0,4142 \quad (\text{koefisien korelasi})$$

$$F = \frac{1}{\frac{(1-0,1716)}{(15-1-1)}} = \frac{0,1716}{0,0637} = 2,694$$

Kesimpulan: tidak signifikan pada taraf signifikansi 0,05.

Regresi sederhana dari Y atas X_2 :

$$b = \frac{\sum X_2 Y}{\sum X_2^2} = \frac{9,96}{50,4} = 0,197619 \quad a = \bar{Y} - b\bar{X}_2 = 1,74 - 0,197619(3,2) = 1,1076Y'$$

$$= 1,1076 + 0,197619X_2$$

$$SSR = \frac{(\sum X_2 Y)^2}{\sum X_2^2} = \frac{(9,96)^2}{50,4} = 1,9682857$$

$$SSE = SST - SSR$$

$$SSE = 3,777 - 1,9682857 = 1,8087143$$

$$R_{y.2}^2 = \frac{SSR}{SST} = \frac{1,9682857}{3,777} = 0,521124 \quad (\text{koefisien determinasi})$$

$$R_{y.2} = \sqrt{R_{y.2}^2} = \sqrt{0,521124} = 0,7218891 \quad (\text{koefisien korelasi})$$

Kesimpulan: signifikan pada taraf signifikansi 0,05.

Dari kedua kesimpulan tersebut dapat dikatakan bahwa variabel besar keluarga X_2 merupakan yang lebih baik untuk pengeluaran untuk bahan makanan Y dibanding variabel penghasilan keluarga X_1 .

8.3 Analisis Linear Berganda dengan Variabel *Dummy*

Pada analisis regresi linear berganda dengan variabel *dummy*, salah satu atau beberapa variabel independennya merupakan variabel *dummy*. Variabel *dummy* atau variabel kategorikal adalah variabel dengan skala nominal, misalnya seperti variabel jenis kelamin, agama, warna kulit, suku bangsa, jenis pekerjaan, organisasi politik, daerah pemukiman, dan sebagainya.

Variabel *dummy* ini menggunakan kode 1 dan kode 0 sebagai penanda hasil pengamatan. Contohnya adalah kode 1 menunjukkan mahasiswa pria dan 0 untuk mahasiswi, atau misalnya angka 1 untuk pengusaha muda anggota HIPMI dan angka 0 untuk pengusaha muda non anggota HIPMI. Modelnya dapat digunakan seperti persamaan regresi linear berganda di sub bab sebelumnya, di mana salah satu atau beberapa variabel *dummy* independennya merupakan variabel *dummy*.

$$Y' = a + B_1X_1 + B_2X_2 + \dots + B_kX_k$$

Dalam analisis regresi linear berganda (termasuk dengan variabel *dummy*) peneliti dapat memasukkan sekian banyak (k) variabel independen. Persamaan regresi linear, uji F, serta uji-t nya dapat diperoleh dengan menggunakan alat bantu statistik.

Dalam analisis ini dapat diilustrasikan perilaku kewirausahaan mahasiswa di fakultas ekonomi universitas X. Untuk itu dilakukan

pengamatan lima variabel independen yang diprediksi berpengaruh terhadap perilaku kewirausahaan mahasiswa fakultas ekonomi universitas X tersebut.

Variabel-variabel penelitiannya adalah:

Y = Skor perilaku kewirausahaan mahasiswa

X₁ = Jenis kelamin

X₂ = Pengalaman kerja apapun (dalam tahun)

D₁ = Kondisi keluarga (*variabel dummy*, keluarga keturunan wirausaha =1, bukan keturunan wirausaha = 0)

D₂ = Asal kampung (*variabel dummy*, kota = 1, desa = 0)

D₃ = Jenis motivasi (*variabel dummy*, uang = 1, non uang = 0)

Melalui alat bantu statistik, diperoleh persamaan regresi linear berganda dengan variabel *dummy* sebagai berikut (n – 36)

$$Y' = 41,56 + 0,61X_1 + 0,04X_2 + 0,61D_1 + 7,73D_2 + 3,7D_3$$

t (0,98) (0,61) (0,24) (3,05) (2,63)

$$R^2 = 0,61$$

$$F = 2,82$$

Dapat dilihat 5 variabel independen secara simultan atau bersama-sama berpengaruh signifikan terhadap skor perilaku kewirausahaan mahasiswa (harga uji F 2,82 lebih besar daripada nilai kritis F = 2,53). Selanjutnya, secara parsial hanya variabel asal daerah D₂ dan variabel motivasi D₃ yang mempunyai pengaruh yang signifikan terhadap skor perilaku kewirausahaan (berturut-turut t = 3,05 dan t = 2,68 lebih besar daripada nilai kritis t = 2,36). Dalam uji, koefisien determinasi R² = 0,61 menampilkan bahwa 61% dari variasi skor perilaku kewirausahaan mahasiswa dijelaskan oleh kelima variabel bebasnya (jenis kelamin, pengalaman, kondisi keluarga, asal daerah, dan motivasi). Dan 39% -nya dijelaskan oleh variabel-variabel independen lainnya yang tidak diamati dalam penelitian tersebut:

SOAL – SOAL LATIHAN

1. Apa yang dimaksud dengan *Simple Regression Analysis* dan bagaimana penggunaannya? Jelaskan!
2. Apa yang *Multiple Regression Analysis* dan bagaimana penggunaannya? Jelaskan!
3. Riset sederhana dilakukan pada mahasiswa fakultas ilmu ekonomi universitas X, untuk mengetahui apakah ada hubungannya antara kebiasaan menonton berita ekonomi dengan kecenderungan perilaku hemat mereka, penskalaan diukur dengan 100 poin. Kuisisioner disebar pada mahasiswa semester 3, dengan sampel 20 orang dengan asumsi melihat televisi 4 jam sehari. Waktu penelitian dilakuka 2 minggu berturut-turut. Dari riset diperoleh data:

Variabel kebiasaan menonton berita ekonomi (X)									
60	30	60	39	50	30	40	54	58	28
50	45	47	34	57	53	55	49	59	34

Variabel kecenderungan perilaku berhemat (Y)									
58	33	69	35	44	49	29	48	54	60
34	45	42	56	42	58	53	76	45	54

Pertanyaannya :

- a Bagaimana bentuk persamaan regresinya?
- g. Jika para mahasiswa menonton 75 jam selama 2 minggu berapa kecenderungan perilaku hemat para mahasiswa?

BAB IX

PROSEDUR PENGUJIAN STATISTIK NONPARAMETRIK

9.1 Statistik Nonparametrik

Dalam praktik riset, ranah kajian untuk ilmu – ilmu non eksakta atau bidang ilmu sosial sering menemukan kesulitan mendapatkan data kontinu yang menyebar mengikuti distribusi normal. Data penelitian ilmu-ilmu sosial yang diperoleh kebanyakan hanya berupa kategori yang hanya dapat dihitung frekuensinya atau berupa data yang dapat dibedakan berdasarkan tingkatan atau rankingnya. Menghadapi kasus data kategorial atau data ordinal seperti itu, jelas peneliti tidak mungkin mempergunakan metode statistik parametrik. Sebagai gantinya diciptakan oleh para pakar metode statistik lain yang sesuai, yaitu yang disebut metode statistik nonparametrik.

Metode statistik nonparametrik ini sering juga disebut metode bebas sebaran (*distribution free methods*) karena model uji statistiknya tidak menetapkan syarat-syarat tertentu tentang bentuk distribusi parameter populasinya. Artinya, bahwa metode statistik nonparametrik ini tidak menetapkan syarat bahwa observasi-observasinya harus ditarik dari populasi yang berdistribusi normal dan tidak menetapkan syarat *homoscedascicity*. Sejumlah uji statistik nonparametrik hanya menetapkan asumsi/persyaratan, bahwa observasi-observasinya harus independen dan variabel yang diteliti pada dasarnya harus memiliki kontinuitas. Banyak di antara uji-uji statistik nonparametrik terkadang disebut sebagai “uji ranking” karena teknik-teknik nonparametrik ini dapat digunakan untuk skor yang bukan skor eksak dalam pengertian angka, melainkan berupa skor yang semata-mata berupa jenjang-jenjang (ranks).

Hasil pemikiran para pakar untuk menciptakan metode-metode statistik nonparametrik, ternyata dapat menunjukkan hasil yang cukup baik, tidak jauh berbeda dengan hasil yang diperoleh dengan metode statistik parametrik. Metode statistik nonparametrik mempunyai kelebihan-kelebihan bila dibandingkan dengan metode statistik parametrik, selain kekurangan-kekurangannya.

Manfaat atau kelebihan metode statistik nonparametrik dibanding metode statistik parametrik, antara lain dapat disebutkan sebagai berikut:

1. Nilai probabilitas dari sebagian besar uji statistik nonparametrik diperoleh dalam bentuk yang lebih pasti (kecuali untuk kasus sampel yang besar), tak peduli bagaimana bentuk distribusi populasi yang merupakan induk dari sampel-sampelnya. Ketepatan nilai probabilitas itu tidaklah bergantung pada bentuk distribusi populasinya, meskipun beberapa uji statistik nonparametrik mungkin menganggap bahwa distribusi populasinya simetris. Dalam kasus-kasus uji nonparametrik tertentu menganggap, bahwa distribusi yang mendasarinya adalah kontinu, suatu anggapan yang juga dibuat dalam uji-uji parametrik.
2. Apabila sampel-sampelnya kecil atau terpaksa kecil karena sifat hakikat sampel itu sendiri (misalnya $n = 6$), hanya uji statistik nonparametrik yang dapat digunakan, kecuali jika sifat distribusi populasinya diketahui secara pasti.
3. Uji-uji statistik nonparametrik dapat digunakan untuk menganalisis data yang pada dasarnya merupakan jenjang atau ranking juga untuk data yang skor-skornya secara sepiantas tampak memiliki kekuatan ranking, dan bagi data yang hanya dapat dikategorikan sebagai plus atau minus, lebih atau kurang, lebih baik atau lebih buruk, dan sebagainya.
4. Uji-uji statistik nonparametrik dapat digunakan untuk menganalisis data yang hanya merupakan klarifikasi semata, yakni data yang terukur dalam skala nominal.
5. Tersedia uji-uji statistik nonparametrik untuk menganalisis sampel-sampel yang terdiri dari observasi-observasi dari beberapa populasi yang berlainan.
6. Uji-uji statistik nonparametrik sederhana perhitungannya sehingga lebih mudah dipelajari dan diterapkan dibanding dengan uji-uji parametrik.

Di samping kelebihan-kelebihan tersebut, uji-uji statistik nonparametrik juga mempunyai kekurangan-kekurangan. Adapun kekurangan-kekurangan yang dapat dikemukakan dari uji statistik nonparametrik adalah:

1. Apabila persyaratan-persyaratan bagi model statistik parametrik (terutama asumsi distribusi normal) dapat dipenuhi dan apabila pengukuran data mempunyai kekuatan seperti yang disyaratkan, pemakaian uji statistik nonparametrik, kekuatan efisiensinya menjadi lebih rendah.

2. Uji statistik nonparametrik tidak dapat dipergunakan untuk menguji interaksi seperti dalam model analisis variance.
3. Metode statistik nonparametrik tidak dapat dipergunakan untuk membuat prediksi (ramalan) seperti dalam model analisis regresi karena asumsi distribusi normal tidak dapat dipenuhi.
4. Macam uji statistik nonparametrik terlalu banyak sehingga menyulitkan peneliti dalam memilih uji mana yang paling sesuai.

9.2 Beberapa Jenis Uji Statistik Nonparametrik

Uji statistik nonparametrik macamnya banyak atau boleh dibilang macamnya banyak sekali. Menurut Siegel (1956) uji statistik nonparametrik ada 27 buah. Karena macamnya yang banyak sekali, maka tidak mengherankan apabila terkadang peneliti menemui kesulitan dalam memilih uji statistik nonparametrik mana yang paling cocok dengan kasus penelitiannya.

Dalam memilih macam uji statistik nonparametrik mana yang paling sesuai, pertama-tama kita mendasarkan pada tingkatan skala pengukuran variabel penelitiannya, apakah berskala nominal, ordinal, atau berskala interval. Pada dasarnya, uji yang sesuai bagi variabel dengan skala nominal atau skala ordinal adalah uji statistik nonparametrik, namun terdapat juga uji nonparametrik yang berlaku bagi variabel dengan skala interval. Kedua, pemilihan didasarkan pada banyaknya sampel penelitian, apakah berupa sampel tunggal (satu sampel), dua sampel berpasangan atau dua sampel independen, atau sampelnya lebih dari dua buah yang berpasangan atau yang independen. Ketiga, pemilihan uji nonparametrik didasarkan pada jenis penelitiannya, apakah berupa uji kesesuaian (*goodness-of-fit*), berupa uji banding, atau berupa uji keterikatan (korelasi) antara dua kumpulan atribut atau antara dua kumpulan skor.

Skala Pengukuran	Kasus Sampel Tunggal	Kasus Dua Sampel		Kasus K Sampel		Pengukuran Korelasi Nonparametrik
		Sampel-sampel Berpasangan	Sampel-sampel Independen	Sampel-sampel Berpasangan	Sampel-sampel Independen	
Nominal	Uji Binomial Uji Chi-square sampel Tunggal	Uji MnNemar untuk Signifikansi Perubahan	Uji Kemungkinan Fisher Uji Chi-Square Dua Sampel Independen	Uji Cochran		Koefisien kontingensi C

Ordinal	Uji kolmogoro v-smimov sampel Tunggal Uji Runs Sampel Tunggal	Uji tanda uji Berjenjang wilcoxon untuk Data Berpasangan	Analisis Ragam dua arah Friedman	Uji medium lanjutan Analisis Ragam satu arah Kruskal-walls	Koefisien kontingensi Rank Spearman Koefisien korelasi Kendal Koefisien relasi Rank Parsial Kendall Koefisien Konordansi Kendall
Interval/Rasio		Uji Walsh uji Randomisasi untuk Data Berpasangan			

Misalnya, jika penelitian bertujuan untuk menguji *goodness-of-fit*, sampelnya hanya satu dan pengukuran variabelnya berskala nominal (berupa kategori), maka dapat digunakan uji Chi-Square sampel tunggal. Apabila berupa uji banding sampelnya dua buah independen, dan pengukuran variabelnya berskala ordinal atau ranking, maka dapat digunakan uji U dari Mann-Whitney, dan apabila peneliti lainnya bertujuan untuk mengukur korelasi antara dua kumpulan skor, di mana skor-skor tersebut berupa skala ordinal, maka dapat digunakan analisis korelasi rank Spearman atau analisis korelasi rank Kendall .

Dalam Bab X ini tidak semua macam uji Statistik nonparametrik (27 buah) dibicarakan, tetapi hanya akan dipilih beberapa macam Uji statistik nonparametrik yang dipertimbangkan sering dibicarakan dalam buku-buku statistik pada umumnya. Beberapa macam uji Statistik nonparametrik tersebut adalah:

1. Uji Binomial (termasuk uji tentang proporsi populasi)
2. Uji jenjang Bertanda Wilcoxon (untuk dua sampel berpasangan);
3. Uji Walsh;
4. Uji Chi- Square untuk dua sampel ndependen (termasuk uji beda dua proporsi):
5. Uji U dan Mann-Whitney :
6. Uji Kolmogorov- Smirnov untuk dua sampel independen
7. Analisis variance dua arah dari Friedman;
8. Ui Chi- Square untuk banyak sampel independen;
9. Analisis variance satu arah dari Kruskall Walls;
10. Koefisien Kontingensi C; dan

11. Analisis Korelasi Rank Spearman.

9.3 Uji Binomial

Uji Binomial ini berlaku untuk kasus sampel tunggal di mana datanya berskala nominal dengan dua atribut. Terdapat populasi - populasi yang dipandang hanya terdiri dari dua kelompok. Misalnya, kelompok lelaki dan perempuan (variabel jenis kelamin), rusak dan tidak rusak (produksi barang), anggota dan bukan anggota (keanggotaan koperasi), sekolah dan tidak sekolah (pendidikan), dan sebagainya.

Untuk populasi apa saja yang hanya terdiri dari dua kelompok, jika kita melambangkan proporsi kasus-kasus yang masuk dalam kelompok yang satu dengan p , maka proporsi dari kelompok yang satunya adalah $1 - p$ (biasanya dilambangkan dengan q).

Besarnya proporsi p dalam satu kelompok dan proporsi q dalam kelompok lainnya, dari hasil pengamatan pada sampel akan berbeda dengan proporsi p dan proporsi q dari populasinya sebagai akibat fluktuasi random dalam Pengambilan sampel penelitiannya. Tujuan dan pengujian adalah untuk menyimpulkan apakah perbedaan tersebut hanya karena faktor kebetulan semata atau perbedaan tersebut benar-benar signifikan,

Uji binomial didasarkan pada distribusi kemungkinan binomial. Distribusi kemungkinan binomial adalah distribusi sampling dari proporsi-proporsi yang mungkin kita amati dalam sampel sampel random yang ditarik dari suatu populasi yang terdiri dari dua kelompok.

Probabilitas untuk memperoleh X obyek dalam satu kategori/kelompok dan $(n - X)$ obyek dalam kategori lainnya, dihitung dengan rumus:

$$P(X) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

Pengujian hipotes dapat didasarkan pada sampel kecil ($n \leq 25$) atau didasarkan pada sampel besar ($n \geq 25$). Dalam kasus satu sampel kecil di mana besarnya $P = q$, probabilitas yang berkaitan dengan harga-harga sekecil harga-harga X observasi sudah tersusun dalam tabel (Lihat Lampiran Tabel E) jadi kita tertolong dari perhitungan satu per satu dengan menggunakan rumus distribusi binomial tersebut. Lampiran Tabel E menyajikan kemungkinan-kemungkinan (probabilitas) satu sisi berkaitan dengan

terjadinya bermacam-macam harga yang seekstrem X di bawah hipotesis nihil bahwa $p = q = 0,5$. Apabila kasusnya $p \neq q$ kita terpaksa menggunakan rumus distribusi binomial tersebut.

Untuk kasus satu sampel besar, Lampiran Tabel E tidak dapat kita gunakan, Sebagai gantinya kita menggunakan pendekatan kurva normal dengan rumus-rumus sebagai berikut:

Mean	$\mu = n.p$
Deviasi standar	$\sigma = \sqrt{n.p.q}$
Hatga uji statistik	$Z = \frac{(x \pm 0,5) - \mu}{\sigma}$
	$Z = \frac{(x \pm 0,5) - n.p}{\sqrt{n.p.q}}$

Catatan:

Angka 0,5 merupakan *adjustment* (dari diskrit menjadi kontinu) Untuk $X < n.p$ tambahkan 0,5 pada X, dan bila $X > n.p$ kurangkan 0,5 pada X.

Langkah-langkah dalam menggunakan uji binomial:

1. Tentukan n, yaitu banyaknya kasus yang akan diobservasi (banyaknya sampel)!
2. Hitunglah jumlah frekuensi pada masing-masing kategori.
3. Metode menentukan probabilitas terjadinya suatu harga, atau harga yang lebih ekstrem di bawah H_0 .
 - a. Apabila $n \leq 25$, dan jika $p = q = 0,5$. Lampiran Tabel E menyajikan probabilitas satu sisi tentang munculnya berbagai harga yang sekecil harga X observasi, di bawah H_0 . Untuk uji dua sisi, harga p kalikan dengan dua.
 - b. Apabila $p \neq q$, probabilitas akan terjadinya harga X, atau yang lebih ekstrem, di bawah H_0 , dihitung dengan cara mensubstitusikan harga-harga observasi dalam rumus distribusi probabilitas binomial. Perhitungan ini dapat dibantu dengan Lampiran Tabel F₂ (Tabel Koefisien Binomial).
 - c. Apabila $n > 25$ dan p sama atau mendekati 0,5, ujliah H_0 dengan menggunakan pendekatan distribusi normal.

Contoh Sampel Kecil :

Dalam kajian psikologis dampak kelelahan, seorang ahli psikologi dan juga seorang peneliti, mencontohkan pada 18 anggota pramuka dari

UKM Universitas X , mengenai teknik simpul dengan tali yang sama, namun tekniknya berbeda. Separuh dari subjek tersebut (dipilih secara random dari 18 anggota pramuka), belajar dengan teknik A lebih dahulu, dan separuh lainnya belajar dengan teknik B lebih dahulu. Selanjutnya, dalam keadaan kelelahan (setelah dilakukan kegiatan seharian) setiap subjek diminta menyimpulkan tali dengan teknik semula.

Pendugaannya ialah keadaan lelah dan tertekan simpul yang termemori adalah cara pertama. Selanjutnya subjek digolongkan menurut cara mereka menyimpul dalam keadaan lelah dan tertekan.

Teknik membuat simpul tali yang dipilih dalam keadaan stress adalah:

	Teknik yang Dipilih		Total
	Yang Pertama Dipelajari	Yang Kedua Dipelajari	
Frekuensi	16	2	18

Proses pengujiannya dilakukan sebagai berikut:

1. Formulasi hipotesis nihil (H_0) dan hipotesis alternatif (H_1):

$H_0 : P_1 = P_2 = 0,5$ Artinya, tidak ada perbedaan antara kemungkinan (probabilitas) menggunakan teknik pertama yang dipelajari di bawah tekanan dan kelelahan (P_1), dan kemungkinan menggunakan teknik kedua yang dipelajari di bawah tertekan dan lelah (P_2).

$H_1 : P_1 > P_2$ = Artinya, kemungkinan menggunakan teknik pertama yang dipelajari di bawah tekanan dan kelelahan (P_1), lebih besar daripada kemungkinan menggunakan teknik kedua yang dipelajari di bawah tekanan dan kelelahan (P_2).

Uji yang sesuai adalah uji binomial.

2. Taraf signifikansi dan berapa sampel.

Digunakan taraf signifikansi $\alpha = 0,01$ dengan $n = 18$.

3. Daerah penolakan

Daerah penolakan terdiri dari semua harga X (banyaknya subjek yang menggunakan teknik kedua yang diajarkan, dalam keadaan lelah) sedemikian kecilnya sehingga kemungkinan yang berkaitan dengan kejadian di bawah H_0 adalah sama atau lebih dari $\alpha = 0,01$ ($P \leq 0,01$). Karena arah perbedaannya sebelumnya telah diramalkan maka daerah penolakannya terletak pada satu sisi (pengujian satu arah).

Kriteria pengujiannya adalah:

Bila Harga Uji Statistik $P < 0,01$ H_0 ditolak

$P > 0,01$ H_0 diterima

4. Harga uji statistik

Karena sampelnya kecil ($n = 18$) dan $P = q = 0,5$ maka Lampiran Tabel E dapat digunakan. Untuk $n = 18$, kemungkinan yang berkaitan dengan $X \leq 2$ adalah $p = 0,001$.

5. Kesimpulan atau keputusan pengujian

Karena harga uji statistik $p = 0,001$ lebih kecil daripada $\alpha = 0,001$, maka keputusannya H_0 ditolak dan ini berarti H_1 diterima pada taraf signifikansi 0,01. Disimpulkan bahwa $P_1 > P_2$ yakni bahwa murid-murid yang berada di bawah stress, akan kembali ke teknik yang dipelajari pertama di antara dua teknik yang ada.

Contoh Sampel Besar

Masih dengan ilustrasi diatas, namun dengan 30 anggota pramuka, diperesetasikan data sebagai berikut :

	Teknik yang Dipilih		Total
	Yang Pertama Dipelajari	Yang Kedua Dipelajari	
Frekuensi	26	4	30

Proses pengujian seperti contoh sampel kecil:

$H_0 : P_1 = P_2 = 0,5$

$H_1 : P_1 > P_2$ atau $P_2 < P_1$

1. Hipotesis
2. Digunakan uji Binomial dengan pendekatan kurva normal.
3. Digunakan taraf signifikansi 0,01 dengan $n = 30$.
4. Daerah penolakan. Nilai kritis pada pengujian satu sisi kiri: $Z_{0,01} = 2,33$. Lihat Lampiran Tabel Kurva Normal!

Kriteria pengujiannya:

Apabila harga uji statistik: $Z < -2,33$ H_0 ditolak

$Z > -2,33$ H_1 diterima

5. Harga uji statistik Z:

$$Z = \frac{(x \pm 0,5) - n.p}{\sqrt{n.p.q}}$$
$$Z = \frac{(4+0,5) - (0,5)(30)}{\sqrt{n.p.q}} = -3,834$$

6. Keputusan:

Karena harga uji statistik $Z = -3,834$ lebih kecil daripada harga/nilai kritis $Z_{0,01} = -2,33$ maka keputusannya H_0 ditolak dan H_1 diterima, pada taraf signifikansi 0,01. Disimpulkan bahwa $P_1 > P_2$ atau $P_2 < P_1$

9.4 Uji tentang Proporsi Populasi

Apabila $p \neq q$, sampel penelitiannya harus diperbesar. Untuk sampel yang berukuran besar kita dapat menggunakan pendekatan kurva normal. Uji proporsi ini didasarkan atas distribusi sampling harga proporsi dengan *mean* $\mu = \pi$ dan $\sigma_p = \sqrt{\pi(n - \pi)/n}$

Contoh Uji Proporsi:

Data prediktif sebuah media kota Medan dalam pemberitaanya menampilkan suatu informasi bahwa 60% dari warga kota tersebut berprofesi sebagai pedagang (sehingga dapat dikatakan sejumlah 40% bukan pedagang) atau $P (\pi) = 60\%$ dan $Q = 40\%$). Data tersebut masih spekulatif karena belum diuji secara statistik. Untuk keperluan itu suatu kajian dilakukan dengan 200 orang ditarik dengan acak sebagai sampel dan mereka adalah pedagang.

Proses pengujiannya:

1. Hipotesis: $H_0 : = 0,60$
 $H_1 : < 0,60$
2. Digunakan uji proporsi (uji Z)
3. Digunakan taraf signifikansi 0,05 dengan $n = 200$
4. Harga/nilai kritis, pengujian satu sisi kiri, nilai $Z_{0,05} = -1,645$.
5. Harga uji statistik:

$$Z = \frac{\hat{p} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$
$$Z = \frac{0,50 - 0,60}{\sqrt{\frac{0,60(1-0,60)}{200}}} = \frac{-0,10}{0,0346} = -2,89$$

Atau dihitung Dengan rumus :

$$Z = \frac{X - n.p}{\sqrt{n.p.q}}$$

$$Z = \frac{100 - 120}{\sqrt{(100)(0,06)(0,40)}} = \frac{-20}{6,928} = -2,89$$

6. Keputusan:

Karena harga uji statistik $Z = -2,89$ lebih kecil daripada harga/nilai kritis $Z_{0,05} = -1,645$. maka H_0 ditolak dan berarti H_1 diterima ($\pi < 0,60$, proporsi pedagang kurang dari 60%) pada taraf signifikansi 0,05.

9.5 Uji Jenjang Bertanda Wilcoxon

Mekanisme uji ini berlaku untuk dua buah sampel yang berpasangan dengan skor yang berskala ordinal. Uji ini pada dasarnya merupakan penyempurnaan dari “Uji Tanda”, yakni selain memperhatikan tanda positif dan negatif, besarnya beda juga tak luput dari perhatian. Asumsi penting dalam uji ini adalah bahwa masing-masing perbedaan antara pasangan skor (d_i) merupakan distribusi yang sifatnya simetris,

Dugaan nol, atau hipotesis nihil (H_0) berbunyi bahwa dua populasinya bersifat identik. Andai H_0 benar ada harapan bahwa jumlah jenjang yang bertanda positif kira-kira akan seimbang dengan jumlah jenjang yang bertanda negatif. Bila dua jumlah jenjang itu beda jauh diantara yang satu dengan yang lain konklusinya adalah bahwa dua populasi itu tidak identik sehingga H_0 ditolak. Dalam kalimat lain, H_0 ditolak bila salah satu jumlah jenjang positif atau negatif sangat kecil.

Pengujian jenjang Wilcoxon dapat didasarkan pada sampel kecil ($n \leq 25$) atau sampel besar ($n \geq 25$). Untuk sampel kecil, pengujian didasarkan pada nilai T, yakni jumlah yang lebih kecil antara jumlah jenjang positif dengan jumlah jenjang negatif. Lampiran Tabel menunjukkan berbagai nilai T dengan taraf signifikansi 0,05, 0,02, dan 0,01 untuk pengujian satu dan dua sisi. Sementara untuk sampel besar, diuji dengan distribusi normal, di mana *mean* dan deviasi standar dari distribusi sampling nilai T dihitung dengan rumus:

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

$$\text{Deviasi standar: } \sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$$\text{Harga Uji Statistik } Z = \frac{T - U_T}{\sigma_t}$$

$$Z = \frac{T - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}$$

Pentahapan dalam uji jenjang Wilcoxon dapat ditempuh sebagai berikut:

1. Untuk masing-masing pasangan skor dihitung beda atau selisihnya (d_i). Beda ini bisa bersifat positif dan bisa bersifat negatif.
2. Tahapan berikutnya adalah memberikan penjenjangan harga d_i dengan tidak memperhatikan tandanya, dimulai dari yang terkecil hingga yang terbesar. Andai ada harga-harga d_i yang sama maka perlu dilakukan penghitungan jenjang rata-ratanya!
3. Selanjutnya, memberikan tanda nilai (bisa tanda positif atau negatif) pada jenjang untuk tiap-tiap beda sesuai dengan tanda dari beda itu tersebut. Untuk beda 0 diabaikan, andai terdapat beda 0, kembali n, yaitu banyak total harga d_i yang memiliki tanda.
4. Kemudian, menetapkan besaran nilai T, yaitu jumlah yang lebih kecil dibanding dua kelompok jenjang yang memiliki tanda yang sama, positif atau negatif.
5. Terakhir, bahwa prosedur yang digunakan dalam menetapkan signifikansi harga T sampel, bergantung pada besarnya n:
 - a. Apabila $n \leq 25$, Lampiran Tabel G menyajikan harga-harga T untuk berbagai ukuran
 - b. $n (n \leq 25)$. Jika harga T observasi < nilai T tabel maka H_0 ditolak.
 - c. Apabila $n > 25$, H_0 diuji dengan menggunakan pendekatan kurva normal.

Contoh untuk Sampel Kecil

Riset pasar perlu dilakukan untuk melihat apakah ada perbedaan kepuasan distributor (distributor satisfaction) pada pra dan pasca implementasi kebijakan pelayanan terbaru pada perusahaan produsen *magicom* (penanak nasi) di Surabaya. Adapaun, data yang berhasil dihimpun adalah sebagai berikut :

Agen	Skor Sebelum	Skor Sesudah
1	82	72
2	32	39

3	44	48
4	49	50
5	73	64
6	20	25
7	24	30
8	52	50
9	39	32
10	39	35
11	22	33
12	42	48
13	17	25
14	71	66
15	58	60
16	41	47
17	56	59
18	15	27
19	36	40
20	55	58

Adapun secara teknis dapat dilakukan proses pengujian yang alurnya sebagai berikut :

1. Formulasi H_0 dan H_1
 H_0 : Tidak terdapat perbedaan kepuasan para agen antara sebelum dan sesudah adanya kebijaksanaan pelayanan yang baru.
 H_1 : Terdapat perbedaan kepuasan para agen antara sebelum dan sesudah adanya kebijaksanaan pelayanan yang baru.
2. Uji statistik yang sesuai adalah uji T dari Frank Wilcoxon.
3. Taraf signifikansi dan besar sampel.
 Misalnya digunakan $\alpha = 0,05$. Nilai T untuk uji jenjang bertanda Wilcoxon adalah (untuk $n = 20$) $T_{0,05} = 52$.
4. Kriteria pengujian:
 Apabila harga uji statistik: $T < 52$ H_0 ditolak
 $T > 52$ H_0 diterima
 Catatan: T adalah jumlah jenjang yang lebih kecil.
5. Perhitungan harga uji statistik T seperti pada tabel berikut:
 Tabel 10.2

Komputasi Harga Uji Statistik kepuasan Agen, Sehubungan Kebijakan Baru

Agen	Skor Sebelum	Skor Sesudah	Beda Skor	Jenjang Gabungan	+d	-d
1	82	72	-10	18		18
2	32	39	7	15	15	
3	44	48	4	8	8	
4	49	50	1	1	1	
5	73	64	-9	17		17
6	20	25	5	10,5	10,5	
7	24	30	6	13	13	
8	52	50	-2	2,5		2,5
9	29	32	3	5	5	
10	39	35	-4	8		8
11	22	33	11	19	19	
12	42	48	6	13	13	
13	17	25	8	16	16	
14	71	66	-5	10,5		10,5
15	58	60	2	2,5	2,5	
16	41	47	6	13	13	
17	56	59	3	5	5	
18	15	27	12	20	20	
19	36	40	4	8	8	
20	55	58	3	5	5	
T= 56						

6. Kesimpulan dari pengujian.

Karena harga uji statistik $T = 56$ lebih besar daripada harga/nilai kritis $T_{0,025=52}$ maka diputuskan H_0 diterima pada taraf signifikansi 0,05. Disimpulkan, bahwa data penelitian tidak cukup mendukung perubahan kepuasan para agen berkaitan dengan adanya perubahan kebijakan pelayanan.

Contoh Sampel Besar :

Dilakukan penelitian untuk mengetahui apakah terdapat perbedaan kecepatan Yang dihitung dalam jumlah kata per kata per menit pengetikan antara model mesin ketik lama dan model mesin ketik baru. Dari percobaan dengan 29 orang pengetik diperoleh hasil sebagai berikut :

Pengetik	Mesin Lama	Mesin Baru
1	43	49
2	91	92
3	33	32
4	54	54
5	45	65
6	55	90
7	65	64
8	90	85
9	53	56
10	70	70
11	76	74
12	87	87
13	32	64
14	99	104
15	87	87

Pengetik	Mesin Lama	Mesin Baru
16	80	72
17	88	88
18	23	32
19	75	90
20	54	51
21	43	49
22	23	90
23	56	78
24	56	57
25	70	70
26	76	78
27	45	60

28	76	80
29	54	54

Prosedur pengujiannya dilakukan sebagai berikut:

1. Formulasi H_0 dan H_1

H_0 : Jumlah Jenjang Positif seimbang dengan jumlah jenjang negatif.
Artinya, mesin ketik model baru tidak berpengaruh terhadap kecepatan dalam pengetikan yang diukur dalam jumlah kata per menit.

H_1 : Mesin ketik model baru berpengaruh terhadap kecepatan dalam pengetikan yang diukur dalam jumlah kata per menit.

2. Digunakan Uji T dari Frank Wilcoxon.

3. Taraf signifikansi dan besar sampel.

Misalnya digunakan $\alpha = 0,05$. Dalam kasus tersebut $n = 29$ (sampel besar dengan $n > 25$), maka digunakan pendekatan kurva normal. Lihat Lampiran Tabel A! Karena arah perbedaannya sudah diramalkann, maka digunakan pengujian satu sisi. Harga/nilai kritis $Z_{0,05} = 1,645$.

4. Kriteria pengujian:

Apabila harga uji statistik: $Z < 1,645$ H_0 ditolak

$Z > 1,645$ H_0 diterima

5. Perhitungan harga uji statistik T seperti pada tabel berikut:

Pengetik	Mesin Lama	Mesin Baru	Beda (d)	Jenjang d	Tanda -	Jenjang-
1	43	49	6	13,5	13,5	
2	91	92	1	2,5	2,5	
3	33	32	-1	2,5		2,5
4	54	54	0			
5	45	65	20	18	18	
6	55	90	35	21	21	
7	65	64	-1	2,5		2,5
8	90	85	-5	11,5		11,5
9	53	56	3	8	8	
10	70	70	0			
11	76	74	-2	5,5		5,5

12	87	87	0		
13	32	64	32	20	20
14	99	104	5	11,5	11,5
15	87	87	0		
16	80	72	-3	8	
17	88	88	0		
18	23	32	9	15	
19	75	90	15	16,5	16,5
20	54	51	-3	8	8
21	43	49	6	13,5	13,5
22	23	90	67	22	22
23	56	78	22	19	19
24	56	57	1	2,5	2,5
25	70	70	0		
26	76	78	2	5,5	5,5
27	45	60	15	16,5	16,5
28	76	80	4	10	10
29	54	54	0		
				215	30

Catatan:

Beda 0 tidak diperhatikan (ada 7 buah).Jadi $n = 29 - 7 = 22$

$$\text{Mean; } \mu_1 = \frac{n(n+1)}{4} = \frac{22(22+1)}{4} = 126,5$$

$$\text{Deviasi standar : } \sigma_1 = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$$\sigma_1 = \sqrt{\frac{22(22+1)(44+1)}{24}}$$

$$\sigma_1 = \sqrt{\frac{(506)(45)}{24}} = 30,802$$

$$\text{Harga Uji Statistik: } Z = \frac{T - \mu_1}{T}$$

$$Z = \frac{30 - 126,5}{30,802} = - 2,87 \text{ (sisi kiri)}$$

$$Z = \frac{215 - 126,5}{30,802} = - 2,87 \text{ (sisi kanan)}$$

6. Kesimpulan :

Karena harga uji statistik $Z = 2,87$ lebih besar daripada harga/nilai kritis $Z_{0,05} = 1,645$ atau $Z_{0,05} = -2,87$ lebih kecil dari $Z_{0,05} = -1,645$ maka

diputuskan H_1 ditolak pada taraf signifikansi 0,05. Dapat disimpulkan bahwa ketik model baru berpengaruh nyata terhadap peningkatan kecepatan dalam pengetikan yang diukur dalam jumlah kata per menit.

9.6 Uji Walsh

Pengujian model Walsh ini berlaku untuk menguji dua sampel berpasangan dengan corak data yang berskala interval. Asumsi/anggapan dari uji ini bahwa skor-skor selisih (d_i) yang amati dalam dua sampel yang berpasangan berasal dari populasi simetris dan juga tidak populasi yang sama. Jika populasinya simetris, maka nilai *mean* merupakan gambaran yang akurat dari nilai tengah, dan sama dengan nilai median.

Beda/selisih masing-masing n pasangan (d_i) diurutkan berdasarkan besarnya. Dimulai dengan d_1 = skor beda yang terendah (mungkin negatif), d_2 = skor beda kedua dari yang terendah, dan seterusnya. Dengan demikian $d_1 \leq d_2 \leq d_3 \dots \leq d_n$. Hipotesis nihil yang akan diuji mengatakan, bahwa harga-harga d_i itu ditarik dari suatu populasi yang nilai mediannya = 0 (atau dari sekelompok populasi yang memiliki :”median bersama” sama dengan nol). Uji walsh menganggap bahwa sejumlah d_i itu adalah dari populasi dengan distribusi simetris. Dalam distribusi simetris, nilai mean dan median berimpit, maka H_0 mengatakan bahwa $\mu_1 \neq 0$ (uji dua sisi) atau $\mu_1 > 0$ atau $\mu_1 < 0$ (uji satu sisi).

Selanjutnya guna menentukan signifikansi, berbagai hasil dengan uji Walsh dapat dilihat dalam Tabel Nilai Kritis, yang menampilkan harga signifikansi untuk uji satu dan dua sisi. Pentahapan langkahnya adalah sebagai berikut :

1. Menentukan skor beda dalam tanda (d_i) untuk tiap pasangan yang. Dijodohkan!
2. Menentukan n yakni banyaknya pasangan yang dijodohkan!
3. Urutkan harga-harga d_i berdasarkan besarnya (dari d_1 hingga d_n)!Tanda heda diperhatikan sehingga d_1 adalah d negatif yang terbesar dan d_n adalah positif yang terbesar .
4. Nilai Kritis Uji Walsh Lampiran Tabel H digunakan untuk menentukan apakah H_0 ditolak dan menerima H_1 dengan harga-hargad $d_1, d_2, d_3, \dots, d_n$ yang diobservasi

Contoh Uji Walsh (Siegel, 1956):

Misalkan dilakukan kajian psikologis mengenai memori, yang dilakukan dalam keadaan tertekan / ditekan. Sampel 15 mahasiswa fakultas ekonomi universitas X yang ditugasi untuk menghafal 10 kata acak. Peneliti memberikan intervensi yang sifatnya mengganggu yakni dengan membentak ketika responden menyebut 5 kata tertentu yang dipilih acak dari 10 kata tersebut. Selanjutnya selang dua (2) hari kemudian, responden di minta menyebutkan kembali kata – kata tersebut. Opini peneliti adalah para mahasiswa akan cenderung ingatan kata – kata yang ketika penyebutan tidak disertai bentakan sebagai pengganggu.

Jumlah kata yang disertai gangguan dan tidak ketika penyebutan, setelah 2 hari berselang.

Subjek	Diingat Tanpa Gangguan	Diingat Dengan Gangguan
A	5	2
B	4	2
C	3	0
D	5	3
E	2	3
F	4	2
G	2	3
H	2	1
I	4	1
J	4	3
K	3	4
L	1	2
M	5	2
N	4	4
O	1	0

Adapun arah dan tahapan ujinya sebagai berikut :

1. Hapotesis

H_0 : Nilai tengah / median beda antara jumlah kata tanpa gangguan yang diingat dan banyak kata yang disertai gangguan adalah nol, Artinya, subjek akan mengingat kata itu sama baiknya ($\mu_0 = 0$)

H_1 : Jumlah kata yang diingat tanpa gangguan lebih banyak daripada yang disertai gangguan ($\mu_0 > 0$)

2. Digunakan uji Walsh.
3. Taraf signifikansi dan besar sampel,
Misalnya digunakan $\alpha = 0,05$. Dalam kasus ini besar sampel $n = 15$.
4. Kriteria pengujian:
Karena H_1 adalah $\mu_1 > 0$ maka H_0 akan ditolak jika ada di antara harga-harga yang dicantumkan dalam kolom sebelah kanan dari Lampiran Tabel Hitung, untuk $n = 15$, ternyata terjadi. Alasannya adalah bahwa taraf signifikansi untuk semua harga yang ditabelkan untuk $n = 15$ adalah lebih kecil daripada $\alpha = 0,05$.
5. Perhitungan harga uji statistik Walsh:

Jumlah kata yang disertai gangguan dan tanpa gangguan berhasil diingat oleh setiap subyek sesudah 2 hari berlalu, beserta beda (d) untuk masing-masing subjek disajikan dalam tabel berikut.

Subjek	Diingat Tanpa Gangguan	Diingat dengan Gangguan	Beda Skor d
A	5	2	3
B	4	2	2
C	3	0	3
D	5	3	2
E	2	3	-1
F	4	2	2
G	2	3	-1
H	2	1	1
I	4	1	3
J	4	3	1
K	3	4	-1
L	1	2	-1
M	5	2	3
N	4	4	0
O	1	0	1

Perhatikan bahwa d yang terkecil -1 , dengan demikian $d_1 = d$ terendah, dengan memperhatikan tandanya. Lima di antara harga-harga d adalah -1 ; karenanya $d_1 = -1$, $d_2 = -1$, $d_3 = -1$, $d_4 = -1$, $d_5 = -1$. Harga d yang terkecil setelah itu adalah 1 . Tiga subjek (H , J , dan O) memiliki d yang besarnya 1 Maka $d_6 = 1$, $d_7 = 1$, $d_8 = 1$. Nilai d berikutnya 2 , ada tiga. Dengan demikian $d_9 = 2$, $d_{10} = 2$, dan $d_{11} = 2$. Harga d yang terbesar adalah 3 . Banyaknya 4 . Maka $d_{12} = 3$, $d_{13} = 3$, $d_{14} = 3$ dan $d_{15} = 3$. Berdasarkan Lampiran Tabel H ditunjukkan bahwa untuk $n = 15$ pengujian satu sisi bagi H_1 bahwa $\mu > 0$ pada $\alpha = 0,047$ adalah: Minimum $[1/2(-d_1 + d_{12}), 1/2(d_2 + d_{11})] > 0$

Kata “minimum berarti kita harus memilih yang lebih kecil di antara dua harga itu, dalam kaitannya dengan harga d yang ada di dalam soal tersebut. Artinya, jika $1/2(-d_1 + d_{12})$ atau $1/2(d_2 + d_{11})$, dipilih mana yang lebih kecil, ternyata lebih besar dari nol, maka kita dapat menolak H_0 pada taraf signifikansi $0,047$. Dari tabel tersebut, diketahui bahwa $d_1 = -1$ dan $d_{12} = 3$, $d_2 = -1$ dan $d_{11} = 2$, dengan mensubstitusikan harga-harga itu, kita dapatkan:

$$\text{Minimum } [1/2(-1 + 3), 1/2(-1 + 2)]$$

$$\text{Minimum } [1/2(2), 1/2(1)] = 1/2(1) = 1/2$$

6. Keputusannya dari kedua harga tersebut yang lebih kecil adalah $4\% = 1$. Karena harga d lebih besar dari nol maka H_0 ditolak pada $\alpha = 0,047$. Kita simpulkan bahwa jumlah kata yang berhasil diingat tidak disertai gangguan nyata lebih besar daripada jumlah kata yang diingat oleh subjek yang disertai gangguan.

9.7 Uji *Chi-Square*- Dua Sampel Independen

Pengujian dalam Uji *chi-square* yang akan dibicarakan disini adalah uji *chi-square* yang berlaku untuk kasus dua sampel independen di mana datanya berskala nominal (kategorial). Uji *chi square* ini digunakan untuk menguji signifikan perbedaan antara dua kelompok yang independen. Hipotesis yang akan diuji mengatakan bahwa kedua kelompok itu berbeda dalam hal karakteristik tertentu, dengan demikian perbedaan itu berhubungan dengan frekuensi relatif masuknya anggota-anggota kelompok (dua kelompok) ke dalam beberapa kategori (dua atau lebih). Untuk menguji hipotesis ini, kita menghitung banyak kasus dan masing masing kelompok

(dua kelompok) yang termasuk dalam berbagai kategori (dua atau lebih) dan membandingkan proporsi kasus-kasus dari satu kelompok dalam berbagai kategori dengan proporsi kasus-kasus dari kelompok yang lain.

Untuk kasus dua kelompok dengan dua kategori (Tabel Kontingensi 2 x 2), nilai *chi-square* dihitung dengan rumus:

$$X^2 = \frac{n[(AD-BC) - \frac{n}{2}]^2}{(A+B)(D+D)(A+C)(B+D)}$$

Nilai-nilai kritis *chi-square* dengan berbagai taraf signifikansi dan derajat bebas (db) = 1 dapat dilihat pada Lampiran Tabel C untuk - Nilai X_2 .

Untuk kasus dua kelompok dengan kategori lebih dan dua (Tabel Kontingensir x k), nilai *ch-square* dihitung dengan nimus:

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

$\sum_{i=1}^r$ = jumlah semua baris (r)

$\sum_{j=1}^k$ = jumlah semua kolom (k)

O_{ij}	Frekuensi pengamatan (observasi) dari baris ke-I pada kolom ke-j
E_{ij}	Frekuensi diharapkan (teoretis) dari baris ke-I pada kolom ke-j, E_{ij} diperoleh dengan: $(n_1)(n_2)/n$.

Harga-harga X_2 yang diperoleh dari rumus tersebut mengikuti distribusi chi-kuadrat dengan db = (r - 1) (k - 1) di mana r = banyaknya baris dan k = banyaknya kolom dalam tabel kontingensi.

Contoh Tabel Kontingensi 2x2 (Siegel 1982)

Dalam masa pandemi covid 19 ini, kebutuhan akan vaksinasi menjadi penting, karena vaksinasi akan mengurangi kemungkinan seseorang terpapar virus tersebut. Dengan kata lain orang yang telah menerima vaksin covid 19 akan lebih kebal dari pada yang tidak. Dalam ilustrasi ini terdapat dua kelompok, X dan Y yang masing masing terdiri dari 100 . Vaksin covid 19 diberikan kepada kelompok A sedangkan kelompok B tidak diberi vaksin tersebut (kelompok kontrol), kemudian didapati bahwa 75 orang dari kelompok A dan 65 orang dari kelompok B (kelompok kontrol) tidak terserang / terpapar virus covid 19. Gambarnya dapat diperhatikan tabel berikut ini!

	Kelompok A (diberi vaksin)	Kelompok B (tidak diberi vaksin)	Jumlah
Tidak Terpapar	75	65	140
Terpapar	25	35	60
	100	100	100

Prosedur pengujiannya adalah:

1) Formulas H_0 dan H_1

H_0 : Jumlah jenjang positif seimbang dengan jumlah jenjang negatif. Artinya, pemberian vaksin covid 19 tidak memberikan kekebalan mutlak terhadap kemungkinan terpapar virus covid 19.

H_1 : Pemberian vaksin covid 19 memberikan kekebalan mutlak terhadap kemungkinan terpapar virus covid 19.

2) Uji statistik yang sesuai adalah uji X^2 .

3) Taraf signifikansi dan besar sampel.

Misalnya digunakan $\alpha = 0,05$ Dalam kasus tersebut $n_A = 100$ dan $n_B = 100$.

4) Kriteria pengujian:

Berdasarkan Lampiran Tabel C, harga/nilai kritis $X^2_{0,05;db_1} = 3,84$

Apabila harga uji statistik: $X^2 > 3,84$ H_0 ditolak

$X^2 < 3,84$ H_0 diterima

1. Perhitungan harga uji statistik X^2 :

Perhatikan Tabel 10. 5!

	Kelompok A (diberi serum)	Kelompok B (tidak diberi serum)	Jumlah
Sembuh	75	65	140
Tidak sembuh	25	35	60
	100	100	100

$$X^2 = \frac{n \left[(AD - BC) - \frac{n}{2} \right]^2}{(A+B)(D+D)(A+C)(B+D)}$$

$$X^2 = \frac{200 \left[(75)(35) - (65)(25) - \frac{200}{2} \right]^2}{(75+65)(25+35)(75+25)(65+35)}$$

$$X^2 = \frac{200(810.000)}{(84.000.000)} = 1,93$$

2. Kesimpulan/keputusan pengujian:

Karena harga uji statistik $X^2 = 1,93$ lebih kecil daripada harga/nilai kritis $\bar{X}^2_{0,05;db_1} = 3,84$ maka diputuskan H_0 diterima pada taraf signifikansi 0,05. Dapat disimpulkan bahwa tidak ada hubungan mutlak antara pemberian vaksin dengan kemungkinan kebal dari paparan virus covid 19.

Apabila digunakan uji Z (uji beda dua proporsi), proses pengujiannya adalah:

1) Rumusan hipotesis

$H_0 : \pi_a = \pi_b$ Artinya, proporsi sembuh dari kelompok A (diberi vaksin) sama dengan proporsi sembuh dari kelompok B (tanpa vaksin)

$H_1 : \pi_a \neq \pi_b$ Artinya, proporsi sembuh dari kelompok A (diberi vaksin) tidak sama dengan proporsi sembuh dari kelompok B (tanpa vaksin).

2) Digunakan uji Z

3) Digunakan $\alpha = 0,05$ dengan besar sampel $n_A = 100$ dan $n_B = 100$.

4) Didasarkan pada distribusi sampling harga beda dua proporsi yang mengikuti distribusi normal. Perhatikan Lampiran Tabel A! Harga/nilai kritis $Z_{0,25} = \pm 1,96$ (pengujian dua sisi).

5) Harga uji statistik Z:

$$Z = \frac{X_A/n_A - X_B/n_B}{\sqrt{\frac{\pi_A(1-\pi_A)}{n_A} + \frac{\pi_B(1-\pi_B)}{n_B}}}$$

Hipotesis nihil mengatakan bahwa $\pi_a = \pi_b$ atau $\pi_a - \pi_b = 0$, maka:

$$Z = \frac{\frac{X_A}{n_A} - \frac{X_B}{n_B}}{\sqrt{\pi(1-\pi)(1/n_A + 1/n_B)}}$$

π ditaksir dengan proporsi gabungan dari dua sampel independen yang diambil.

$$\pi = \frac{X_A}{n_A} + \frac{X_B}{n_B}$$

$$\pi = \frac{76}{100} + \frac{65}{100} = \frac{140}{100} = 0,70$$

$$Z = \frac{75/100 - 65/100}{\sqrt{0,70(1 - 0,70)(1/100+1/100)}}$$

$$Z = \frac{0,10}{\sqrt{(0,21)(0,02)}} = \frac{0,10}{0,0648} = 1,54$$

3. Kesimpulan/keputusan:

Karena harga uji statistik $Z = 1,54$ berada di antara harga/nilai kritis $Z_{0,25} = -1,96$ dan $+1,96$ maka diputuskan H_0 diterima pada taraf signifikansi $0,05$. Disimpulkan bahwa $Z_a = Z_b$, artinya bahwa proporsi orang yang sembuh dari kelompok A (yang diberi vaksin sama saja dengan proporsi orang yang sembuh dari kelompok B (tidak diberi vaksin), ini berarti serum tersebut tidak memberi kekebalan mutlak atas kemungkinan terpapar covid 19.

Contoh Tabel Kontingensi $r \times k$:

Dari kawasan berikat industri di kota Medan, diambil sampel sebanyak 420 orang pegawai dari berbagai perusahaan swasta yang ada. Sampel tersebut dibedakan menurut jenis kelaminnya sebagai pria dan wanita. Peneliti ingin memeriksa apakah variabel jenis kelamin ada hubungannya dengan hobi para pegawai tersebut berolah raga. Hasil wawancara dengan mereka menghasilkan informasi sebagai berikut:

Sikap	Jenis Kelamin		Jumlah
	Pria	Wanita	
Senang	128	85	213
Tidak Senang	66	78	144
Tidak Menjawab	26	37	63
Jumlah	220	200	420

Proses pengujiannya dilakukan sebagai berikut:

1. Rumusan hipotesis

H_0 : Variabel jenis kelamin tidak ada hubungannya dengan sikap pegawai terhadap hobi berolah raga.

H_1 : Terdapat hubungan antara jenis kelamin dengan sikap pegawai terhadap hobi berolah raga.

2. Digunakan Uji X^2 .

3. Digunakan taraf signifikansi $0,05$ dengan besar sampel $n = 420$ (dengan n pria = 220 dan n wanita = 200).

4. Kriteria pengujian :

Perhatikan Lampiran Tabel C! Untuk tabel kontingensi 3x2, derajat bebas $(3-1)(2-1) = 2$, Nilai /harga kritis $X^2_{0,05 \text{ db } 2} = 5,99$.

Apabila harga uji statistik: $X^2 > 5,99 H_0$ ditolak

$X^2 < 5,99 H_0$ diterima

5. Perhitungan harga uji statistik X^2 , perhatikan tabel berikut!

Komputasi Harga Uji Statistik X^2

Sikap	Jenis Kelamin		Jumlah
	Pria	Wanita	
Senang	126(111,6)	85(101,4)	213
Tidak Senang	66(74,4)	78(68,6)	144
Tidak Menjawab	26(33,0)	37(30,0)	63
Jumlah	220	200	420

Catatan :

$$E_{11} = (213)(220)/420 = 111,6$$

$$E_{12} = (213)(200)/420 = 101,4$$

$$E_{21} = (144)(220)/420 = 75,4$$

$$E_{22} = (144)(200)/420 = 68,6$$

$$E_{31} = (63)(220)/420 = 33,0$$

$$E_{32} = (63)(200)/420 = 30,0$$

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

$$\frac{(126 - 111,6)^2}{111,6} + \frac{(85 - 101,4)^2}{101,4} + \frac{(66 - 75,4)^2}{75,4} + \frac{(78 - 68,6)^2}{68,6}$$

$$+ \frac{(26 - 33)^2}{33} + \frac{(37 - 30)^2}{30}$$

$$X^2 = 2,41 + 2,65 + 1,17 + 1,29 + 1,48 + 1,63 = 10,63$$

6. Kesimpulan/keputusan pengujian:

Karena harga uji statistik $X^2 = 10,63$ lebih besar daripada harga/nilai kritis $X^2_{0,05 \text{ db } 2} = 5,99$ maka diputuskan H_0 ditolak dan H_1 diterima pada taraf signifikansi 0,05. Dapat disimpulkan, bahwa terdapat hubungan antara jenis kelamin dengan sikap terhadap hobi berolah raga.

Uji chi square mensyaratkan bahwa frekuensi yang diharapkan E_i dalam masing-masing sel tidak boleh terlampau kecil (kurang dari 5). Untuk tabel kontingensi 2x2, penggunaan *ujichi square*, disarankan: (1) Bila $n > 40$ gunakan X^2 dengan koreksi kontinuitas (gunakan rumus untuk tabel kontingensi 2 x 2), (2) Bila n berada di antara 20 dan 40, uji X^2 dengan rumus untuk tabel kontingensi 2 x 2 boleh digunakan bila semua frekuensi diharapkan adalah 5 atau lebih. Jika frekuensi diharapkan yang terkecil kurang dari 5 pakailah uji Fisher (periksa contoh di bawah), dan (3) Bila $n < 20$, pakailah uji Fisher untuk kasus apapun.

Untuk tabel kontingensi $r \times k$, berarti d lebih besar daripada 1, uji X^2 dapat digunakan jika kurang dari 20% di antara sel-sel itu mempunyai frekuensi diharapkan yang kurang dari 5 dan jika tidak ada satu sel pun memiliki frekuensi diharapkan yang kurang dari satu. Jika persyaratan tersebut tidak terpenuhi, peneliti dapat menggabungkan kategori kategori yang berdekatan agar frekuensinya dapat diperbesar.

Contoh Uji Kemungkinan Eksak Fisher:

Dalam dunia pemasaran, sebenarnya kaum pria dan wanita sama – sama berhak untuk menjadi tenaga pemasaran. Ilustrasi ini ingin melihat apakah terdapat perbedaan yang nyata antara *sales counter* pria dan wanita pada sebuah dealer mobil dalam menjual produk tersebut. Data penelitian ditunjukkan sebagai berikut:

Pria	Wanita
47	35
50	40
55	35
30	43
40	40
47	
50	

Prosedur pengujiannya dilakukan sebagai berikut:

1. Rumusan hipotesis.

H_0 : Tidak terdapat perbedaan antara *sales counter* pria dan wanita dalam keberhasilan menjual produk.

H₁: Terdapat perbedaan yang nyata antara *sales counter* pria dan wanita dalam keberhasilan menjual produk.

2. Digunakan uji kemungkinan eksak Fisher.
3. Digunakan taraf signifikansi 0,05 dengan besar sampel n = 12.
4. Kriteria pengujian:

Apabila harga $P < 0.05$ H₀ statistik ditolak.

5. Menghitung harga $P > 0,05$ H₀ uji statistik p:

Masing-masing kelompok dihitung mediannya. Jika volume penjualannya sama atau lebih besar dari median dianggap berhasil (+), bila dibawah median dikatakan gagal (-). Median *sales counter* pria = 47 dan median *sales counter* wanita = 40. Perhatikan tabel berikut!

	Kriteria		
	(+)	(-)	
Pria	5(A)	2(B)	7
Wanita	1(C)	4(D)	5
Jumlah	6	6	12

Harga p dihitung dengan rumus:

$$P = \frac{((A+B)!(C+D)!(A+C)!(B+D)!}{n!A!B!C!D!}$$

$$P = \frac{7!5!6!6!}{12!5!2!1!4!}$$

$$P = \frac{(5.040)(120)(720)(720)}{(479.001.600)(120)(2)(24)} = 0,1136$$

6. Kesimpulan / keputusan pengujian:

Karena harga uji statistik $p = 0,1136$ lebih besar dari 0,05 maka diputuskan H₀ diterima pada taraf signifikansi 0,05. Dapat disimpulkan bahwa tidak terdapat perbedaan antara *sales counter* pria dan wanita dalam keberhasilan menjual suatu produk.

9.8 Uji Maan-Whitney

Uji Mann-Whitney disebut juga sebagai uji U, berlaku untuk kasus dua sampel independen dengan skor yang berskala ordinal. Uji Mann-Whitney dipakai Untuk menguji apakah dua kelompok independen telah ditarik dari populasi yang sama. Uji ini merupakan pengembangan dari Wilcoxon dengan dua sampel yang berukuran tidak sama, dan pemberian jenjang didasarkan pada skor gabungan.

Uji Mann-Whitney tidak memerlukan anggapan tertentu mengenai populasi dari mana sampel diambil (seperti uji-uji nonparametrik yang lain). Asumsi yang diperlukan hanyalah bahwa nilai dari variabel random dari dua kelompok yang diperbandingkan adalah berdistribusi kontinu. Hipotesis nihil yang akan diuji mengatakan, bahwa dua sampel independen diambil dari populasi-populasi yang mempunyai distribusi yang sama. Uji ini dapat digunakan untuk pengujian dua sisi ataupun satu sisi. Uji tersebut merupakan alternatif lain dari uji t parametrik, bila anggapan yang diperlukan bagi uji t tidak dijumpai.

Bila besar sampel dinyatakan dengan n_1 (yang lebih kecil) dan n_2 (yang lebih besar), uji Mann-whitney dilakukan sebagai berikut:

1. Untuk sampel yang amat kecil (n_2 tidak lebih dari 8), digunakan Tabel I yang memuat nilai-nilai U dalam Uji Mann-whitney.
2. Untuk sampel yang berukuran lebih besar (n_2) yakni antara 9 dan 20, uji Mann-whitney menggunakan Lampiran Tabel J yang menyajikan nilai-nilai kritis U untuk tarag signifikansi 0,001; 0,01; 0,025; dan 0,05 untuk uji satu sisi atau 0,002; 0,002; 0,05 dan 0,10 untuk uji dua sisi.
3. Untuk sampel besar ($n_2 > 20$), uji Mann-whitney dilakukan berdasarkan pendekatan kurva normal, dengan:

$$\text{Mean} = \mu_1 = \frac{n_1 n_2}{2}$$

$$\text{Deviasi standar: } \sigma_1 = \sqrt{\frac{(n_1)(n_2)(n_1+n_2+1)}{12}}$$

$$\text{Harga uji statistik: } Z = \frac{\mu - \mu_1}{\sigma_1}$$

$$Z = \frac{\mu - n_1 n_2 / 1}{\sqrt{(n_1)(n_2)(n_1+n_2+1)/12}}$$

Tahapan dalam penggunaan uji Mann-Whitney diikhtisarkan sebagai berikut:

1. Tahapan pertama menentukan skor/harga dari n_1 dan n_2 , di mana n_1 = banyaknya pengamatan dari kelompok yang lebih kecil dan n_2 = banyaknya pengamatan dari kelompok yang lebih besar!
2. Selanjutnya adalah menggabungkan kedua sampel independen dan beri jenjang pada tiap-tiap skornya mulai dari skor terkecil sampai dengan skor terbesar! Apabila ada dua atau lebih skor yang sama, digunakan jenjang rata-rata.

3. Hitunglah jumlah jenjang masing-masing kelompok dan notasikan dengan R_1 dan R_2 !

4. Kemudian dihitung harga U dengan rumus:

Dari sampel pertama (n_1):

$$U = n_1 + n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

Dari sampel kedua (n_2)

$$U = n_1 + n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

1. Dari dua nilai/harga tersebut yang digunakan adalah harga U yang lebih kecil. Nilai yang lebih besar ditandai dengan U' . Untuk memeriksa apakah yang didapatkan itu harga U atau U' , perlu dibandingkan dengan $n_1n_2/2$. Bila harganya lebih besar dari $n_1n_2/2$ berarti harga tersebut adalah U' . Harga U diperoleh dengan rumus: $U = n_1n_2 - U'$
2. Penetapan signifikansi harga observasi bergantung pada besarnya:
 - a. Bila $n_2 \leq 8$, digunakan Lampiran Tabel I.
 - b. Bila n_2 antara 9 dan 20, digunakan Lampiran Tabel J.
 - c. Bila $n_2 > 20$, gunakan Lampiran Tabel A (Tabel Kurva Normal).

Contoh untuk Sampel Amat Kecil ($n_2 \leq 8$):

Dalam industri berbasis produk yang sulit di jual misalnya, atau bukan *daily consumption* keberadaan tenaga sales menjadi strategis. Sales menjadi ujung tombak perusahaan. Dalam ilustrasi ini akan dilihat apakah kelompok *counter sales* yang telah menjalani diklat penjualan mempeunyai performa sales yang lebih baik daripada yang belum menjalani diklat penjualan. Tiap-tiap kelompok diambil sampel 5 dan 4 orang. Prestasi penjualan yang mereka peroleh adalah:

Kelompok Eksperimen	Kelompok Kontrol
39	55
32	35
37	28
23	26
41	

Proses pengujiannya dilakukan sebagai berikut:

1. Rumusan hipotesis

H_0 : Prestasi penjualan sales counter yang telah menjalani diklat sama dengan prestasi sales counter yang belum menjalani diklat penjualan.

H_1 : : Prestasi penjualan sales counter yang telah menjalani diklat lebih baik dibanding dnegan prestasi sales counter yang belum menjalani diklat penjualan.

2. Digunakan uji U dari Mann-Whitney.

3. Digunakan taraf signikansi 0,05 dengan ukuran sampel $n_2 = 5$ dan $n_1 = 4$.

4. Kriteria pengujian:

Apabila diperoleh harga uji statistik $p < 0,05$ maka H_0 ditolak, bila sebaliknya $p > 0,05$ maka H_1 diterima.

5. Perhitungan harga uji statistik p, perhatikan tabel berikut!

Kelompok Eksperimen	Jenjang	Kelompok Kontrol	Jenjang
39	7	55	9
32	4	35	5
37	6	28	3
23	1	26	2
41	8		
$R_1=26$		$R_2 = 19$	

Dari sampel yang lebih kecil (n_1):

$$U = n_1 + n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U = 4(5) + \frac{4(4+1)}{2} - 19$$

$$U = 20+10-19 = 11$$

Dari sampel yang lebih besar (n_2)

$$U = n_1 + n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U = 4(5) + \frac{5(5+1)}{2} - 26$$

$$U = 20+15-26 = 9 \text{ (diambil yang kecil)}$$

Berdasarkan Lampiran Tabel I, untuk $n_2 = 5$ dengan $n_1 = 4$ dan $U \leq 9$ nilai p adalah 0,452.

6. Kesimpulan/keputusan:

Karena nilai $p = 0,452$ lebih besar dari $\alpha = 0,05$ maka diputuskan H_0 diterima, pada taraf signifikansi 0,05. Dapat disimpulkan, bahwa prestasi penjualan *sales counter* yang telah menjalani diklat penjualan sama saja dengan yang belum menjalani diklat.

Contoh untuk Sampel $9 \leq n_2 \leq 20$:

Kini kebutuhan literasi media dan kebutuhan membaca menjadi semakin penting dalam masyarakat kontemporer ini. Ilustrasi ini akan membandingkan perihal “kegemaran membaca” antara kelompok masyarakat yang belum ada Perpustakaan Desa (kelompok A) dengan kelompok masyarakat lain, yang sudah ada Perpustakaan Desa di kampungnya (kelompok B). Dari kelompok A diambil sampel 10 orang warga ($n_1 = 10$) dan dari kelompok B diambil sampel 18 orang warga ($n_2 = 18$). Diperoleh data sebagai berikut (sudah diurutkan):

Kelompok A	Skor Kegemaran membaca	Kelompok B	Skor Kegemaran membaca
1	6	1	10
2	5	2	9
3	5	3	8
4	4	4	8
5	4	5	8
6	4	6	7
7	4	7	7
8	3	8	7
9	2	9	6
10	2	10	6
		11	6
		12	5
		13	5
		14	5
		15	5
		16	4
		17	3
		18	2

Melalui uji Mann-Whitney ingin diketahui apakah adanya perpustakaan desa dapat meningkatkan kegemaran membaca warga masyarakat desa.

1. Prosedur pengujianya dilakukan sebagai berikut:

H_0 : Kedua kelompok masyarakat itu sama saja tingkat kegemaran membacanya.

H_1 : Kelompok B mempunyai tingkat kegemaran membaca lebih tinggi dibanding kelompok A (digunakan uji satu sisi).

2. Digunakan uji U dari Mann-Whitney.

3. Digunakan α sebesar 0,05 dengan $n_1 = 10$ dan $n_2 = 18$.

4. Kriteria pengujian:

Harga/nilai kritis U untuk uji satu sisi $\alpha = 0,05$ atau untuk uji dua sisi pada $\alpha = 0,10$ dengan $n_1 = 10$ dan $n_2 = 18$ adalah 55 (Lihat Lampiran Tabel J).

H_0 ditolak apabila harga uji statistik $U < 55$, dan

H_1 diterima apabila harga uji statistik $U > 55$. H

5. Harga uji statistik U perhatikan tabel berikut!

Kelompok A	Skor Kegemaran membaca	Jenjang	Kelompok B	Skor Kegemaran Membaca	Jenjang
1	6	18,5	1	10	28
2	5	13,5	2	9	27
3	5	13,5	3	8	25
4	4	8	4	8	25
5	4	8	5	8	25
6	4	8	6	7	22
7	4	8	7	7	22
8	3	4,5	8	7	22
9	2	2	9	6	18,5
10	2	2	10	6	18,5
			11	6	18,5
			12	5	13,5
			13	5	13,5
			14	5	13,5
			15	5	13,5
			16	4	8
			17	3	4,5

	18	2	2
	$R_1=86$		$R_2=320.5$

Dari sampel pertama:

$$\begin{aligned}
 U &= n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \\
 &= 10(18) + \frac{10(10+1)}{2} - 86 \\
 &= 180 + 55 - 86 = 149
 \end{aligned}$$

Dari sampel kedua:

$$\begin{aligned}
 U &= n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \\
 &= 10(18) + \frac{18(18+1)}{2} - 320,5 \\
 &= 180 + 171 - 320,5 = 30,5
 \end{aligned}$$

Diambil nilai U yang kecil, jadi harga uji statistik $U = 30,5$

6. Kesimpulan/keputusan pengujian:

Karena harga uji statistik $U = 30,5$ lebih kecil dari harga/nilai kritis $U = 55$ maka diputuskan H_0 ditolak dan H_1 diterima pada taraf signifikansi 0,05. Dapat disimpulkan, bahwa kelompok masyarakat yang sudah ada perpustakaan desanya, mempunyai kegemaran membaca lebih tinggi dibandingkan dengan kelompok masyarakat yang belum ada perpustakaan desa.

Contoh untuk Sampel Besar ($N_2 > 20$):

Ingatan sejarah perjuangan bangsa menjadi hal yang sangat penting dalam sebuah bangsa. Salah satu manifestasi penghormatan kepada para pejuang dan mencintai negeri misalnya adalah kesediaan untuk mengikuti program pendidikan militer para remaja pria pada usia 18 tahun selama sebulan. Diksi militer ditekankan untuk mempunyai jiwa nasionalisme yang tinggi. Dalam ilustrasi ini di dipisahkan dua kelompok remaja tersebut, yakni kelompok A yang terlebih dahulu dipertontonkan film perjuangan dan kelompok B sebagai kelompok kontrol yang tidak dipertontonkan film tersebut.

Dari kelompok B diambil sampel 16 pemuda ($n_1 = 16$) dan dari kelompok A diambil sampel 23 pemuda ($n_2 = 23$). Skor sikap dan minat mengikuti diksar militer beserta jenjangnya ditunjukkan dalam tabel berikut.

Skor Sikap dan Minat bertransmigrasi Beserta jenjangnya Dari kelompok A dan Kelompok B

Kelompok A	Skor Sikap dan Minat	Jenjang	Kelompok B	Skor Sikap dan Minat	Jenjang
1	26	29,5	1	34	39
2	24	24,5	2	32	38
3	24	24,5	3	30	36
4	20	16	4	30	36
5	20	16	5	30	36
6	20	16	6	28	33
7	20	16	7	28	33
8	18	12	8	28	33
9	16	9,5	9	26	29,5
10	16	9,5	10	26	29,5
11	14	5	11	26	29,5
12	14	5	12	24	24,5
13	14	5	13	24	24,5
14	14	5	14	24	24,5
15	14	5	15	24	24,5
16	12	1	16	22	20,5
			17	22	20,5
			18	20	16
			19	20	16
			20	20	16
			21	16	9,5
			22	16	9,5
			23	12	1,5
		$R_1=200$			$R_2=580,5$

Dari sampel pertama (n_1):

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U = 16 (23) + \frac{16(16+1)}{2} - 200$$

$$U = 368 + 136 - 200 = 304$$

Dari sampel kedua (n_2):

$$U = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

$$U = 16(23) + \frac{23(23+1)}{2} - 580$$

$$U = 368 + 276 - 580 = 64$$

Digunakan pendekatan kurva norma:

$$\text{Mean} = \mu_1 = \frac{n_1 n_2}{2} = \frac{16(23)}{2} = 184$$

$$\begin{aligned} \text{Deviasi standar: } \sigma_1 &= \sqrt{\frac{(n_1)(n_2)(n_1+n_2+1)}{12}} \\ &= \sqrt{\frac{(16)(23)(16+23+1)}{12}} \\ &= 35,0238 \end{aligned}$$

Proses pengujiannya dilakukan sebagai berikut:

1. Rumusan hipotesis

H_0 : Tingkatan sikap dan minat mengikuti diksar militer kelompok pemuda tersebut tidak berbeda.

H_1 : Tingkatan sikap dan minat mengikuti diksar militer kelompok pemuda A lebih tinggi dibanding kelompok pemuda B.

2. Digunakan uji Mann-Whitney karena kedua kelompok tersebut independen dan pengukuran sikap dan minat berskala ordinal.

3. Digunakan taraf signifikansi 0,01 dengan besar sampel $n_1 = 16$ dan $n_2 = 23$ (lebih besar dari 20).

4. Daerah penolakan:

Nilai kritis dengan taraf signifikansi 0,01 (uji satu sisi), berdasarkan tabel kurva normal adalah 2,3 (satu sisi kanan) atau -2,33 (satu sisi kiri).

5. Harga uji statistik Z:

$$Z = \frac{U - \mu_1}{\sigma_1}$$

$$Z = \frac{304 - 184}{35,02} = 3,43 \text{ (satu sisi kanan)}$$

Atau dihitung:

$$Z = \frac{64 - 184}{35,02} = -3,43 \text{ (satu sisi kiri)}$$

6. Kesimpulan/keputusan:

Karena harga uji statistik $Z = 3,43$ lebih besar dari harga/nilai kritis $Z_{0,01} = 2,33$ (atau harga uji statistik $= -3,43$ lebih kecil dari harga/nilai kritis $Z = -2,33$), maka diputuskan H_0 ditolak atau H_1 diterima pada taraf signifikansi 0,01.

Dapat disimpulkan, bahwa tingkatan sikap dan minat mengikuti diksar militer kelompok pemuda A lebih tinggi dibanding kelompok pemuda B.

9.9 Uji Kolmogorov-Smirnov

Uji Kolmogorov — Smirnov yang akan dibicarakan di sini berlaku untuk dua sampel independen dengan pengukuran variabel yang berskala ordinal. Uji Kolmogorov-Smirnov digunakan untuk menguji apakah dua sampel independen tersebut berasal dari populasi yang sama atau dari populasi-populasi yang memiliki distribusi yang sama. Pengujian didasarkan dengan memperhatikan kesesuaian antara dua distribusi frekuensi kumulatif yang disusun dari dua sampel independen tersebut. Jika distribusi frekuensi kumulatif dari dua sampel independen tersebut cenderung berdekatan atau sama maka hipotesis nihil yang menyatakan bahwa dua sampel independen diambil dari populasi-populasi yang berdistribusi sama, bisa diterima. Oleh karena itu, uji ini didasarkan pada nilai selisih maksimum dari dua distribusi frekuensi kumulatifnya. Apabila nilai selisih maksimum ini lebih besar daripada nilai selisih yang diharapkan ini berarti bahwa kesenjangan antara dua distribusi tersebut cukup besar sehingga hipotesis nihil ditolak.

Seperti pada uji-uji statistik nonparametrik yang lain, uji Kolmogorov — Smirnov ini tidak memerlukan anggapan/persyaratan tertentu tentang populasi dari mana sampel diambil. Persyaratan yang diperlukan hanyalah bahwa nilai variabel random dari dua kelompok yang diperbandingkan adalah kontinu.

Untuk menguji apakah skor-skor kelompok eksperimen berbeda dengan skor-skor kelompok kontrol digunakan uji dua sisi, sedangkan untuk menguji apakah skor-skor kelompok eksperimen lebih tinggi dari skor-skor kelompok kontrol digunakan uji satu sisi. Sampel-sampel penelitian dapat didasarkan pada sampel kecil (bila $n_1 = n_2$ dan bila baik n_1 maupun $n_2 \leq 40$) atau didasarkan pada sampel besar (bila n_1 dan $n_2 > 40$).

Tahapan – tahapan penggunaan uji Kolmogorov-Smirnov dan sampel independen, diikhtisarkan sebagai berikut:

1. Susunlah masing-masing kelompok skor dalam suatu distribusi frekuensi kumulatif menggunakan interval-interval atau klasifikasi-klasifikasi yang sama untuk kedua kelompok!
2. Hitunglah selisih frekuensi kumulatif dari kedua kelompok tersebut untuk tiap-tiap interval atau klasifikasi yang ada!
3. Tentukan selisih frekuensi kumulatif yang terbesar (maksimum) dan notasikan dengan D! Untuk uji satu sisi, D adalah selisih terbesar dalam arah yang diramalkan.
4. Signifikansi tidaknya nilai D ditentukan berdasarkan ukuran sampel dan sifat hakikat H_1 :
 - a. Bila $n_1 = n_2$ dan besarnya ≤ 40 , digunakan Lampiran Tabel K, yaitu tabel nilai-nilai kritis $K_{1\alpha}$ untuk taraf signifikansi 0,05 dan 0,01 untuk uji dua sisi dan uji satu sisi.
 - b. Bila n_1 dan n_2 keduanya > 40 di mana n_1 dan n_2 tidak perlu sama, digunakan Lampiran Tabel L. Nilai-nilai kritis D dihitung dengan rumus-rumus yang tercantum pada tabel tersebut (uji dua sisi).
 - c. Bila n_1 dan n_2 keduanya > 40 dengan uji satu sisi, digunakan pendekatan Chi-Square dengan rumus:

$$X^2 = a D^2 \frac{n_1 n_2}{n_1 + n_2}$$

Signifikansi harga uji statistik X^2 tersebut ditentukan dengan membandingkan nilai kritis X^2 dengan taraf signifikansi tertentu dan derajat bebas 2 (Lihat Lampiran Tabel C). Rumus tersebut juga digunakan untuk kasus sampel kecil dengan ukuran yang tidak sama.

Contoh untuk Sampel $n_1 = n_2$, dan besarnya < 40

Didalam masyarakat, kecenderungan orang untuk memakai parfum berbeda-beda, baik anak – anak maupun pria dewasa. Ilustrasi kajian ini akan melihat apakah ada perbedaan yang nyata antara pria muda (usia 20-30 tahun) yang telah menikah dengan pria muda lajang terhadap kebutuhan memakai parfum.

Data penelitian ditunjukkan sebagai berikut:

Respons	Pria sudah menikah	Pria lajang
Sempurna	7	5
Bagus	12	11
Cukup	8	9
Kurang	3	5
	30	30

Prosedur pengujiannya dilakukan sebagai berikut:

1. Rumusan hipotesis

H_0 : Tidak ada perbedaan respons antara pria muda yang telah menikah dengan pria lajang tangga terhadap suatu produk parfum.

H_1 : Terdapat perbedaan respons antara pria muda yang telah menikah dengan pria lajang tangga terhadap suatu produk parfum.

2. Digunakan uji Kolmogorov-Smirnov karena kasusnya berupa dua sampel independen dan pengukuran variabelnya berupa klasifikasi-klasifikasi yang berciri ordinal.

3. Digunakan taraf signifikansi 0,01 dengan ukuran sampel $n^1 = n_2 = 30$ (berupa sampel-sampel kecil).

4. Kriteria pengujian:

Berdasarkan Lampiran Tabel K, harga nilai kritis K_D dengan $n = 3$ dan $\alpha=1,01$ adalah 13 (uji dua sisi).

H_0 ditolak apabila harga uji statistik $K_D > 13$.

H_1 diterima apabila harga uji statistik $K_D < 13$.

5. Harga uji statistik K_D dihitung sebagai berikut:

Perhatikan tabel berikut!

Respons	Pria sudah menikah	Pria belum menikah	$S_{n_1}(X)$	$S_{n_2}(X)$	$S_{n_1}(X) - S_{n_2}(X)$
Sempurna	7	5	30/7	5/30	2/30
Bagus	12	11	30/19	30/16	30/3
Cukup	8	9	30/27	30/25	30/2
Kurang	3	5	30/30	30/30	30/0
	30	30			

Dari tabel tersebut diketahui bahwa nilai D maksimum = maksimum $[S_{n_1}(X) - S_{n_1}(X)] = 3/30$. Harga uji statistik $K_D = 3$ (pembilang dari D maksimum).

6. Kesimpulan/keputusan:

Karena harga uji statistik $K_D = 3S_{n_1}(X) - S_{n_1}(X)$ lebih kecil dari harga nilai kritis $K_D = 13$ maka diputuskan H_0 diterima pada taraf signifikansi 0,01. Dapat disimpulkan, bahwa tidak terdapat perbedaan respons yang nyata antara pria muda yang sudah menikah dengan yang belum mengenai penggunaan parfum.

Contoh untuk Sampel-Sampel Besar (n_1 dan $n_2 > 40$):

Karakteristik pegawai tentu berbeda beda. Dalam ilustrasi ini akan di uji apakah terdapat perbedaan yang nyata dalam hal frekuensi tidak masuk kerja selama satu periode tertentu antara kelompok pekerja pria dan kelompok pekerja wanita di lingkungan perusahaan-perusahaan swasta di kota X. Dari kelompok pekerja pria diambil sampel 100 ($n_1 = 100$) dan dari kelompok pekerja wanita diambil sampel 200 ($n_2 = 200$).

Data distribusi frekuensi tidak masuk kerja dan distribusi frekuensi kumulatif beserta selisihnya ditunjukkan dalam tabel berikut.

Distribusi Frekuensi Tidak Masuk Kerja Pekerja Perusahaan Swasta di Kotamadya ABC

Pekerja	Tidak Masuk Kerja (Hari)					Jumlah
	0	1 - 2	3 - 10	11 - 20	≥ 21	
Pria	12	30	18	30	10	100
Wanita	24	50	40	60	26	200

Pekerja	Frekuensi Kumulatif Relatif				
Pria $S_{n_1}(X)$	0,12	0,42	0,6	0,9	1
Wanita $S_{n_2}(X)$	0,12	0,37	0,57	0,87	1
	0	0,05	0,03	0,03	0

D Maksimum = $[-] = 0,05$.

Proses pengujiannya dilakukan sebagai berikut:

1. Rumusan hipotesis

H_0 : Tidak ada perbedaan yang berarti dalam frekuensi tidak masuk kerja antara pekerja pria dan wanita.

H_1 : Terdapat perbedaan yang berarti dalam frekuensi tidak masuk kerja antara pekerja pria dan wanita (uji dua sisi).

2. Digunakan uji D dua sampel independen dari Kolmogorov – Smirnov
3. Digunakan taraf signifikansi 0,01 dengan besar sampel $n_1 = 10$ dan $n_2 = 200$ (termasuk sampel besar dengan n_1 dan $n_2 > 40$).
4. Kriteria pengujian:

Lihat Lampiran Tabel L! Harga Kritis D dalam uji Kolmogorov-Smirnov dengan dua sampel besar untuk pengujiannya dua sisi pada $\alpha 0,01$ adalah:

$$1,63 \sqrt{\frac{n_1+n_2}{n_1 n_2}} = 1,63 \sqrt{\frac{100+200}{(100)(200)}} = 0,20$$

H_0 ditolak apabila harga D maksimum $> 0,20$.

H_1 diterima apabila harga D maksimum $< 0,20$.

5. Harga uji statistik D:

Berdasarkan Lampiran Tabel 10. D maksimum = $[S_{n_1}(X) - S_{n_2}(X)] = 0,05$.

6. Kesimpulan/keputusan pengujian:

Oleh karena harga uji statistik D (maksimum) = 0,05 lebih kecil dari harga kritis D = 0,20 maka diputuskan H_0 diterima pada taraf signifikansi 0,01. Dapat disimpulkan bahwa tidak terdapat perbedaan yang berarti dalam frekuensi tidak masuk kerja antara pekerja pria dan wanita.

Apabila arah kecenderungannya dapat diramalkan, misalnya frekuensi tidak masuk kerja dari pekerja pria lebih tinggi daripada pekerja wanita (uji satu sisi), dapat digunakan pendekatan *chi-square* dengan rumus:

$$X^2 = 4D^2 \frac{n_1 n_2}{n_1 + n_2}$$

Berdasarkan data dari Lampiran Tabel 10.12 tersebut, maka nilai *chi-square*-nya adalah:

$$X^2 = 4(0,05)^2 \frac{(100)(200)}{100+200} = 0,667$$

Harga/nilai kritis = 9,210. Karena harga uji statistik $X^2 = 0,667$ lebih kecil dari harga/nilai kritis = 9,210 maka diputuskan H_0 diterima pada taraf signifikansi 0,01.

9.10 Uji Friedman

Model pengujian Friedman biasa untuk k sampel berpasangan ($k > 2$) dengan data yang berskala sekurang-kurangnya ordinal. Uji ini dipergunakan sebagai alternatif dari teknik analisis variance dua arah apabila uji ini tidak memerlukan persyaratan bahwa populasi dari mana sampel diambil berdistribusi normal dan mempunyai varian yang homogen. Oleh karena itu lengkapnya uji ini.

Hipotesis nihil yang akan diuji mengatakan, bahwa sampel penelitiannya berasal dari populasi yang sama. Untuk uji Friedman, data penelitian disusun ke dalam suatu tabel dua arah yang mempunyai n baris dan k kolom. Baris menunjukkan subjek dan kolom menunjukkan kondisi. Data kondisi untuk masing-masing subjek disusun dalam bentuk jenjang (ranking). Selanjutnya, uji Friedman adalah untuk menentukan apakah jumlah keseluruhan jenjang, dinotasikan dengan R_1 , berbeda secara signifikan atau tidak.

Apabila banyaknya baris dan kolom tidak terlalu kecil maka nilai yang dinyatakan dengan rumus:

$$X_1^2 = \frac{12}{nk(k+1)} \sum_{j=1}^k (R_j)^2 - 3n(k+)$$

akan mengikuti distribusi *chi-square* dengan derajat bebas $k - 1$. Di mana:

n = Banyaknya baris

k = Banyaknya kolom

R_j = Jumlah ranking dalam kolom j

Apabila banyaknya baris dan atau kolom kurang dari minimal, yakni untuk $k = 3$ dan $n = 2$ sampai 9 dan untuk $k = 4$ dan $n = 2$ sampai 4, kita gunakan Lampiran Tabel M, yakni tabel probabilitas yang berkaitan dengan harga/nilai sebesar harga-harga atau nilai observasi dalam analisis varianeranking dua arah Friedman.

Tahapan teknis pengujian model Friedman diikhtisarkan sebagai berikut:

1. Menyusun data penelitian ke dalam suatu tabel dua arah yang mempunyai n baris (subjek) dan k kolom (kondisi)!
2. Pada tiap – tiap baris, berilah jenjang pada skor-skornya mulai dari 1 hingga k.
3. Jenjang dari setiap kolom, kemudian kita jumlahkan, dan notasikan dengan R_1 .

4. Nilai/harga X^2 , dihitung dengan rumus:

$$X^2 = \frac{12}{nk(k+1)} \sum_{j=1}^k (R_j)^2 - 3n(k+)$$

5. Nilai kritis untuk X^2 bergantung pada ukuran n dan k:

- a. Lampiran Tabel M memberikan probabilitas yang eksak yang berkaitan dengan harga observasi untuk $k = 3$ dan $n = 2$ hingga 9 dan untuk $k = 4$ dan $n = 2$ hingga 4.
- b. Untuk n dan atau k yang lebih besar dari yang ditunjukkan dalam Lampiran Tabel M, probabilitas yang berkaitan dapat ditentukan dengan melihat distribusi nilai *chi-square* dengan derajat bebas $k - 1$ (Lihat Lampiran Tabel C).

Contoh Penggunaan Tabel M:

Misalkan kita ingin mempelajari skor-skor 3 kelompok di bawah kondisi ($n = 3$ dan $k = 4$).Tiap-tiap kelompok terdiri dari 4 subjek berpasangan,masing-masing satu subjek dihadapkan pada satu kondisi. Skor-skor yang diperoleh dari studi ini adalah sebagai berikut:

Kelompok	Kondisi			
	I	II	III	IV
A	9	4	1	7
B	6	5	2	8
C	9	1	2	6

Dengan analisis variance jenjang dua arah Friedman, ujilah hipotesis nihil bahwa semua sampel kolom-kolom berasal dari populasi yang sama (Siegel,1956,hlm.166-167)

Proses pengujiannya dilakukan sebagai berikut:

1. H_0 : Semua sampel kolom-kolom berasal dari populasi yang sama
 H_1 : Semua sampel kolom-kolom tidak berasal dari populasi yang sama
2. Digunakan uji Friedman.
3. Digunakan taraf signifikansi 0,05 dengan $n = 3$ dan $k = 4$ (3 baris dan 4 kolom). lihat Lampiran Tabel M!
4. Kriteria pengujian:
 H_0 ditolak apabila $p < 0,05$,
 H_1 diterima apabila $p > 0,05$.
5. Harga uji statistik Friedman:

Jenjang dari 3 kelompok berpasangan di bawah 4 kondisi ditunjukkan pada tabel di bawah ini:

Kelompok	Kondisi			
	I	II	III	IV
A	4	2	1	3
B	3	2	1	4
C	4	1	2	3
	11	5	4	10

6. Kesimpulan / keputusan:

Karena $p = 0,033$ lebih kecil dari $0,05$ maka diputuskan H_0 ditolak. Dapat disimpulkan, bahwa semua sampel kolom tidak berasal dari populasi yang sama.

Contoh Penggunaan Tabel C:

Sektor property yang kian berkembang, mengakibatkan permintaan mesin pendingin atau AC Ruangan semakin tinggi. Berkaitan dengan hal itu dilakukan kajian untuk mengetahui apakah terdapat perbedaan kepuasan para agen selama tiga tahun terakhir sehubungan dengan adanya kebijaksanaan pelayanan yang baru dari perusahaan pembuat produk AC.

Skor kepuasan dari sampel 20 agen ditunjukkan sebagai berikut:

Agen	Skor Kepuasan		
	19A	19B	19C
1	82	76	83
2	32	39	42
3	44	48	46
4	49	50	55
5	73	64	60
6	20	25	31
7	24	30	36
8	52	50	53
9	29	32	38
10	39	35	40
11	22	33	27
12	42	48	40
13	17	25	26
14	71	66	62

15	58	60	64
16	41	47	53
17	56	59	67
18	15	27	36
19	36	40	42
20	55	58	60

Proses pengujian adalah:

1. Rumusan Hipotesis

H_0 : Tidak terdapat perbedaan kepuasan para agen tiga tahun terakhir sehubungan dengan adanya kebijaksanaan pelayanan baru dari perusahaan pembuat produk AC.

H_1 : Terdapat perbedaan kepuasan para agen tiga tahun terakhir sehubungan dengan adanya kebijaksanaan pelayanan baru.

2. Akan digunakan pengujian model Friedman.

3. Diterapkan taraf signifikansi. Berdasarkan contoh menunjukkan bahwa $n = 20$ dan $k = 3$, berarti mengikuti distribusi nilai X_1^2 .

4. Kriteria pengujian:

Berdasarkan tabel *chi-square* (Lampiran Tabel C), nilai kritis. $X_1^2_{0,05 \text{ ab } 3-1} = 5,99$.

H_0 ditolak apabila harga uji statistik $X_1^2 > 5,99$.

H_1 diterima apabila harga uji statistik $X_1^2 < 5,99$.

5. Perhitungan harga uji statistik X_1^2 , seperti pada tabel berikut!

Jenjang Kepuasan Tiga Tahun Terakhir dari 20 Agen Produksi AC.

Agen	Skor Kepuasan		
	19A	19B	19C
1	2	3	1
2	3	2	1
3	3	1	2
4	3	2	1
5	1	2	3
6	3	2	1
7	3	2	1
8	2	3	1
9	3	2	1

10	2	3	1
11	3	1	2
12	2	1	3
13	3	2	1
14	1	2	3
15	3	2	1
16	3	2	1
17	3	2	1
18	3	2	1
19	3	2	1
20	3	2	1
	R ₁ = 52	R ₂ = 40	R ₃ = 28

$$X_1^2 = \frac{12}{20(3)(3+1)} [(52^2 + 40^2 + 28^2)] - 20(3)(4+1)$$

6. Kesimpulan/keputusan:

Karena harga uji statistik $X_1^2 = 14,40$ lebih besar dari nilai kritis $X_{0,05; 2}^2 = 5,99$, maka diputuskan H_0 ditolak dan H_1 diterima pada taraf signifikansi 0,05. Dapat disimpulkan, bahwa data penelitian mengindikasikan adanya perubahan kepuasan para agen produk AC dalam kurun waktu 19A, 19B, dan 19C.

9.11 Uji *Chi-Square* k Sampel Independen ($k > 2$)

Uji *chi-square* k sampel independen ($k > 2$) merupakan perluasan dari uji *chi-square* dua sampel independen. Uji ini berlaku apabila data penelitiannya berupa kategori-kategori yang bersifat diskrit, baik nominal maupun ordinal.

Cara mengoperasikan uji *chi-square* (X^2), pertama-tama kita susun frekuensi-frekuensi dalam kategori-kategori dalam suatu tabel kontingensi $r \times k$ (r menunjukkan baris dan k menunjukkan kolom). Hipotesis nihil yang akan diuji mengatakan, bahwa k sampel frekuensi atau proporsi berasal dari populasi yang sama atau populasi-populasi yang identik. Hipotesis nihil tersebut dapat diuji dengan rumus:

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

Di mana :

$\sum_{i=1}^r$ = jumlah semua baris (r)

$\sum_{j=1}^k$ = jumlah semua kolom (k)

O_{ij} = Frekuensi pengamatan (observasi) dari baris ke - i pada kolom ke-j

E_{ij} = Frekuensi diharapkan (teoretis) dari ke-I pada kolom ke-j

Harga-harga X^2 yang diperoleh dengan rumus tersebut mengikuti *chi-square* dengan derajat bebas $(r - 1)(k - 1)$ di mana r = banyaknya baris dan k = banyaknya kolom dalam tabel kontingensi r x k. Lihat Lampiran Tabel C!

Untuk tabel kontingensi r x k, dengan derajat bebas lebih besar dari 1, uji X^2 dapat digunakan jika kurang dari 20% di antara sel-sel itu mempunyai frekuensi diharapkan (teoretis) yang kurang dari 5 dan jika tidak ada satu sel pun memiliki frekuensi diharapkan yang kurang dari satu. Jika syarat tersebut terpenuhi, peneliti dapat menggabungkan kategori-kategori yang berdekatan agar frekuensinya dapat diperbesar.

Contoh:

Industri UMKM menjadi andalan yang dibutuhkan dalam perekonomian nasional akhir-akhir ini. Dalam konteks itu akan diselenggarakan kajian untuk mengetahui apakah terdapat keterkaitan antara skala usaha UMKM (mikro, kecil, menengah) dengan sikap pelaku usaha UMKM tersebut terhadap regenerasi pelaku usaha UMKM. Diambil sampel acak sebanyak 185 pelaku usaha UMKM diperoleh data sebagai berikut:

Skala Usaha	Sikap Pelaku UMKM					
	a	b	c	d	e	f
Mikro (Tenaga Kerja sampai 19 orang)	12	18	6	30	8	6
Kecil (Tenaga Kerja 20 - 99 orang)	6	10	8	32	6	6
Menengah (Tenaga Kerja 100 orang atau lebih)	5	5	5	12	5	5

Keterangan tentang sikap pelaku UMKM terhadap regenerasi usaha:

A = Harus mengikuti profesi orang tua (sebagai wiraswasta bidang tertentu);

B = Diusahakan bekerja mengikuti profesi orang tua (sebagai wiraswasta bidang tertentu);

C = Tidak mengikuti bidang usaha orang tua asal masih sebagai wiraswasta;

D = Bebas memilih;

E = Diusahakan bekerja di luar profesi orang tua (bukan sebagai wiraswasta);

F = Harus bekerja di luar profesi orang tua (bukan sebagai wiraswasta).

Pengujian hipotesis dilakukan sebagai berikut:

1. Rumusan Hipotesis:

Ho: Tidak ada keterkaitan antara skala usaha dengan sikap pengusaha terhadap regenerasi usaha.

H1: Ada keterkaitan antara skala usaha dengan sikap pengusaha terhadap regenerasi usaha.

2. Digunakan uji X^2 k sampel independen ($k > 2$).

3. Digunakan taraf signifikansi 0,05. Sampel penelitian $n = 185$, tersusun dalam tabel dua arah dengan $r = 3$ dan $k = 6$ (3 baris dan 6 kolom).

4. Kriteria pengujian:

Berdasarkan Lampiran Tabel C, nilai kritis $X^2_{0,05; db(3-1)(6-1)} = X^2_{0,05; db10} = 18,307$.

Hipotesis nihil ditolak apabila harga uji statistik $X^2 > 18,307$, dan diterima apabila harga uji statistik $X^2 < 18,307$.

5. Perhitungan harga uji statistik X^2 seperti pada tabel berikut!

Frekuensi Pengamatan dan Diharapkan Berdasarkan Skala Usaha dan Sikap Pengusaha 185 Pengusaha Swasta

Skala Usaha	Sikap Pengusaha						Jumlah
	a	b	c	d	e	f	
Kecil	12 (10)	18 (14,3)	6 (8,2)	30 (32)	8 (8,2)	6 (7,4)	80
Sedang	6 (5,3)	10 (12,1)	8 (7)	32 (27,2)	6 (7)	6 (6,2)	68
Besar	5 (4,6)	5 (6,6)	5 (3,8)	12 (14,8)	5 (3,8)	5 (3,4)	37
Jumlah	23	33	19	74	19	17	185

Frekuensi diharapkan (teoretis):

$$E_{11} = (80)(23)/185 = 9,9 \quad E_{21} = (68)(23)/185 = 8,4$$

$$E_{12} = (80)(33)/185 = 14,3 \quad E_{22} = (68)(33)/185 = 12,1$$

$$E_{13} = (80)(19)/185 = 8,2 \quad E_{23} = (68)(19)/185 = 7$$

$$E_{14} = (80)(74)/185 = 32 \quad E_{24} = (68)(74)/185 = 27,2$$

$$\begin{aligned}
E_{15} &= (80)(19)/185 = 8,22 & E_{25} &= (68)(19)/185 = 7 \\
E_{16} &= (80)(17)/185 = 7,4 & E_{26} &= (68)(17)/185 = 6,2 \\
E_{31} &= (37)(23)/185 = 4,6 & E_{34} &= (37)(74)/185 = 14,8 \\
E_{32} &= (37)(33)/185 = 6,6 & E_{35} &= (37)(19)/185 = 3,8 \\
E_{33} &= (37)(19)/185 = 3,8 & E_{36} &= (37)(17)/185 = 3,4
\end{aligned}$$

$$\begin{aligned}
X^2 &= \sum_{i=1}^r \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\
&= \frac{(12-10)^2}{10} + \frac{(18-14,3)^2}{14,3} + \frac{(6-8,2)^2}{8,2} + \frac{(30-32)^2}{32} + \frac{(8-8,2)^2}{8,2} + \frac{(6-7,4)^2}{7,4} + \frac{(10-12,1)^2}{8,4} \\
&+ \frac{(8-7)^2}{7} + \frac{(32-27,2)^2}{27,2} + \frac{(6-7)^2}{7} + \frac{(6-6,2)^2}{6,2} + \frac{(5-4,6)^2}{4,6} + \frac{(5-6,6)^2}{6,6} + \frac{(5-3,8)^2}{3,8} \\
&+ \frac{(12-14,8)^2}{14,8} + \frac{(5-3,8)^2}{3,8} + \frac{(5-3,4)^2}{3,4} = 0,400 + 0,957 + 0,590 + 0,125 + 0,05 + 0,265 + 0,68 \\
&6 + 0,036 + 0,143 + 0,847 + 0,143 + 0,006 + 0,035 + 0,388 + 0,379 + 0,753 = 6,667
\end{aligned}$$

6. Keputusan/kesimpulan pengujian:

Karena harga uji statistik $X^2 = 6,667$ lebih kecil dari harga kritis = 18,307 maka diputuskan bahwa H_0 diterima pada taraf signifikansi 0,05 (tidak cukup alasan untuk menolak H_0). Artinya, pada taraf signifikansi 0,05 dapat disimpulkan bahwa memang tidak ada keterkaitan antara skala usaha (kecil, sedang, besar) dengan sikap pengusaha dalam hal regenerasi usaha (a, b, c, d, e, f), baik pengusaha kecil, sedang, maupun besar cenderung memberi kebebasan (d) kepada anak keturunan untuk memilih profesinya masing-masing.

9.12 Uji Kruskal - Wallis

Model pengujian selanjutnya adalah uji Kruskal-Wallis berlaku untuk k sampel bebas/independen ($k > 2$) dengan skor-skor dengan berskala ordinal. Jenis uji Kruskal-Wallis ini merupakan alternatif dari analisis varian satu arah, bila uji ini tidak memerlukan anggapan bahwa populasi dari mana sampel penelitian diambil mempunyai distribusi normal dan mempunyai varian yang sama. Asumsinya adalah bahwa variabel acak di mana berbagai sampel yang dibandingkan berdistribusi kontinu.

Model pengujian ini diarahkan untuk menentukan apakah k sampel bebas ($k > 2$) berasal dari populasi berbeda. Nilai sampel hampir selalu berbeda, namun apakah perbedaan antara nilai-nilai sampel itu menampilkan perbedaan populasi sebenarnya. Hipotesis nol atau hipotesis nihil yang akan

diuji menunjukkan bahwa k sampel berasal dari populasi yang sama atau identik (mempunyai harga rata-rata/*mean* sama). Sementara hipotesis alternatifnya adalah bahwa populasinya memiliki nilai *mean* yang tidak sama, atau setidaknya satu pasang *mean* tidak sama.

Konteks uji ini menampilkan bahwa semua sampel independen (k sampel) digabungkan, lalu semua nilai observasi diberi jenjang dari nilai terkecil sampai nilai terbesar. Jumlah jenjang dari setiap sampel lalu dihitung dan diberi notasi dengan R_j di mana $j = 1, 2, \dots, k$.

Lalu, jika n_j menampilkan ukuran sampel dan ukuran tiap sampel tidak terlalu kecil ($n_j > 5$), maka nilai uji statistik H dirumuskan:

$$H = \frac{12}{n(n+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(n+1)$$

Dimana :

K = Jumlah sampel ($k > 2$)

n_j = Jumlah kasus dalam sampel ke-j

$n = \sum n_j$ = banyak kasus dalam semua sampel

Dalam konteks ini, distribusi nilai H tersebut mendekati distribusi nilai *chi-square* dengan tingkat bebas db = k — 1. Nilai kritis *chi-square* dapat dikonfirmasi pada Lampiran Tabel C. Jika nilai pengamatan H lebih besar dari nilai kritis *chi-square* untuk derajat signifikansi yang telah dipilih, dengan db = k - 1, maka hipotesis nol / nihil ditolak pada derajat signifikansi tersebut.

Jika ukuran setiap sampelnya kecil, yakni untuk k = 3 dan banyak kasus dalam setiap sampelnya lima atau kurang ($n_j \leq 5$), pendekatan dengan distribusi *chi-square* tidak cukup baik. Pada kasus jenis itu, nilai nilai kemungkinannya disajikan dalam Lampiran Tabel N. Tahapan langkah menguji dengan Kruskal-Wallis adalah berikut:

1. Memberikan tingkatan pada harga observasi untuk k kelompok itu dalam suatu urutan dari 1 sampai n!
2. Menentukan nilai R (jumlah jenjang) untuk setiap k kelompok itu!
3. Andai di antara harga dari observasi itu bernilai sama yakni proporsinya besar, harga H dihitung dengan rumus:

$$= \frac{\frac{12}{n(n+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(n+1)}{1 - \frac{7}{n^3 - n}}$$

Keterangannya $T = t^3 - t$ (t adalah banyak observasi berangka sama dalam serangkaian skor berangka sama).

4. Metode untuk menilai signifikansi harga observasi H bergantung pada ukuran k dan pada ukuran masing-masing sampel/kelompok.
 - a. Jika $k = 3$ dan jika $n_1, n_2,$ dan $n_3 < 5$, digunakan Lampiran Tabel N.
 - b. Untuk kasus-kasus lain, digunakan Lampiran Tabel C.

Ilustrasi untuk sampel kecil:

Untuk membandingkan tiga bola bowling, seorang pemain profesional memainkan lima *game* untuk masing-masing bola dan diperoleh hasil sebagai berikut:

Bola Bowling A :	221	232	207	198	212
Bola Bowling B :	202	225	252	218	226
Bola Bowling C :	210	205	189	196	216

Gunakan uji Kruskal-Wallis, pada taraf signifikansi 0,05 untuk menguji hipotesis nihil bahwa hasil permainan bowler tersebut seimbang di antara 3 bola bowling tersebut, (Freund, 1981).

Prosedur pengujiannya adalah:

1. Formulasi H_0 dan H_1

H_0 : Hasil permainan bowler seimbang di antara 3 bola bowling.

H_1 : Hasil permainan bowler tidak seimbang di antara 3 bola bowling.

2. Uji statistik yang sesuai adalah uji varianceranking satu arah Kruskal-Wallis.

3. Digunakan taraf signifikansi 0,05 dengan ukuran masing-masing sampel $n_1 = 5, n_2 = 5$ dan $n_3 = 5$.

4. Kriteria pengujian:

H_0 ditolak Apabila : $p < 0,05$

H_0 diterima Apabila : $p > 0,05$

5. Perhitungan harga uji statistik H :

Hasil permainan Beserta jenjangnya dari Tiga Bola Bowling

Bola Bowling A		Bola Bowling B		Bola Bowling C	
Nilai	Jenjang	Nilai	Jenjang	Nilai	Jenjang
221	11	202	4	210	7

232	14	225	12	205	5
207	6	252	15	189	1
19	3	218	10	196	2
212	8	226	13	216	3
42		54		18	
R ₁		R ₂		R ₃	

$$H = \frac{12}{n(n+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(n+1)$$

$$\begin{aligned}
 H &= \frac{12}{15(15+1)} \left(\frac{42^2}{5} + \frac{54^2}{5} + \frac{18^2}{5} \right) - 3(15+1) \\
 &= 0,05(352,8+583,2+64,8) - 48 \\
 &= 50,04 - 48 = 2,04
 \end{aligned}$$

6. Kesimpulan/keputusan pengujian:

Berdasarkan Lampiran Tabel N, untuk $n_1=5, n_2 =5$ dan $n_3 = 5$, $H > 2,04$ mempunyai kemungkinan muncul, di bawah H_0 sebesar $p > 0,102$. Kesimpulannya, karena $p (> 0,102)$ lebih besar dari $\alpha = 0,05$ maka H_0 diterima. Dapat disimpulkan, bahwa hasil permainan bowler tersebut seimbang di antara 3 bola bowling.

Contoh Sampel Besar:

Sebuah departemen latihan dan pengembangan mengirimkan bahan kursus yang bersifat “Self-instructional” tentang “penyusunan anggaran” kepada tenaga-tenaga penjualan di empat daerah penjualan. Setelah bahan-bahan dikirimkan lengkap, kemudian dilakukan tes standar. Skor yang dicapai ditunjukkan sebagai

Daerah Penjualan	Skor yang dicapai					
A	78	89	73	82	91	
B	60	88	93	54	68	93
C	92	83	87	84	86	90
D	82	70	95	69	78	74

Dengan taraf signifikansi 0,05, ujlah hipotesis nihil yang mengatakan bahwa hasil dari 4 kelompok sampel berasal dari distribusi skor populasi yang sama, (Kazmier, 1979)

Prosedur pengujiannya adalah:

1. Formulasinya H_0 dan H_1
2. Uji statistik yang sesuai adalah uji varian ranking satu arah Kruskal-Wallis.
3. Digunakan taraf signifikansi 0,05 dengan sampel $n_1 = 5, n_2 = 7, n_3 = 6$ dan $n_4 = 6$. Lampiran Tabel C menunjukkan harga kritis $X^2_{0,05;db 3} = 7,81$.
4. Kriteria pengujian:
 - H_0 ditolak Apabila : $H > 7,81$
 - H_0 diterima Apabila : $H < 7,81$
5. Harga uji statistik H dihitung sebagai berikut:

Skor Tes Standar Beserta Jenjangnya dari Empat Daerah Penjualan

A	Jenjang	B	Jenjang	C	Jenjang	D	Jenjang
78	8,5	60	2	92	21	82	10,5
89	18	88	17	83	12	70	5
73	6	93	22,5	87	16	95	24
82	10,5	54	1	84	13	69	4
91	20	68	3	86	15	78	8,5
		93	22,5	90	19	74	7
		85	14				
$R_1 = 63$		$R_2 = 82$		$R_3 = 96$		$R_4 = 59$	

Nilai H (tanpa Korelasi):

$$H = \frac{12}{n(n+1)} \sum_{j=1}^k \frac{R_j^2}{n} - 3(n+1)$$

$$H = \frac{12}{24(24+1)} \left(\frac{63^2}{6} + \frac{82^2}{6} + \frac{96^2}{6} + \frac{59^2}{6} \right) - 3(24+1)$$

$$= 0,02(793,8+960,57+1.536,0+58,67) - 75$$

$$= 0,02(3.870,54) - 75 = 2,41$$

Nilai H (dengan koreksi):

$$t = 2 \quad 2 \quad 2$$

$$T = 6 \quad 6 \quad 6$$

$$1 - \frac{T}{n^2 - n} = 1 - \frac{(6+6+6)}{24^2 - 24} = 0,9987$$

$$H = \frac{2,41}{0,9987} = 2,4131$$

Karena terjadinya skor sama (kembar) tidak banyak maka koreksi terhadap nilai H tidak banyak berarti (dari 2,4 menjadi 2,4131).

6. Kesimpulan/keputusan pengujian:

Karena harga uji statistik H (2,41 dan setelah dikoreksi 2,4131) lebih kecil dari harga kritis $X^2_{0,05;db_3} = 7,81$, maka diputuskan hipotesis nihil diterima pada taraf signifikansi 0,05. Dapat disimpulkan, bahwa empat kelompok sampel berasal dari distribusi skor populasi yang sama.

9.13 Koefisien Kontingensi C

Dalam konteks riset disiplin ilmu-ilmu sosial, umumnya peneliti ingin mengetahui apakah dua himpunan nilai mempunyai korelasi atau tidak. Apabila berkorelasi, lalu berapa besar tingkat hubungan antara keduanya. Dalam kasus statistic parametrik, ukuran korelasi yang biasa digunakan adalah koefisien korelasi *productmoment* dari Karl Pearson (r). Statistik model ini mensyaratkan bahwa nilai-nilai yang digunakan paling tidak tertulis dalam skala interval. Jika peneliti ingin menguji signifikansi harga observasi r, peneliti tidak hanya harus memenuhi persyaratan skala interval itu, namun juga menganggap bahwa skor itu berasal dari populasi normal bivariat. Jika dengan suatu himpunan data tertentu, persyaratan pengukuran untuk r tersebut tidak terpenuhi, maka kita dapat menggunakan salah satu teknik menghitung koefisien korelasi nonparametrik beserta uji signifikansinya.

Koefisien kontingensi C adalah suatu ukuran derajat hubungan, asosiasi, atau dependensi dari klasifikasi-klasifikasi dalam tabel kontingensi r x k (di mana r menunjukkan banyaknya baris dan k menunjukkan banyaknya kolom). Koefisien kontingensi (*coeficientofcontingency*) tersebut dinyatakan dengan rumus:

$$C = \sqrt{\frac{X^2}{X^2 + n}}$$

Di mana:

X^2 = Nilai *chi-square*

n = besar sampel

Tahapan dan langkah untuk menghitung harga C dan uji signifikansinya, diikhtisarkan sebagai berikut:

1. Tahap pertama adalah menyusun frekuensi hasil pengamatan dalam suatu tabel kontingensi rxk!

2. Selanjutnya menghitung frekuensi yang diharapkan/frekuensi teoretis di bawah H_0 untuk masing-masing sel dengan mengalikan kedua jumlah tepi yang sama-sama dimiliki sel itu (jumlah frekuensi baris kali jumlah frekuensi kolom), kemudian hasilnya dibagi dengan n , yakni jumlah total kasus (besar sampel). Apabila lebih dari 20% di antara sel-sel itu mempunyai frekuensi teoretis kurang dari 5, atau apabila ada sembarang sel yang mempunyai frekuensi teoretis kurang dari 1, gabungkanlah kategori-kategori guna meningkatkan jumlah frekuensi teoretisnya. Periksa kembali Bab X sub J tentang uji *chi-square* untuk banyak sampel independen.

3. Dari data tersebut, hitunglah harga *chi-square* dengan rumus:

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

4. Dengan harga X^2 tersebut, hitunglah harga koefisien kontingensi C dengan rumus:

$$C = \sqrt{\frac{X^2}{X^2 + n}}$$

5. Untuk menguji signifikansi dari harga uji statistik C gunakan Lampiran Tabel C (Tabel Harga-harga Kritis dari *Chi-Square*).

Contoh:

Suatu lembaga riset pemasaran mendapatkan order untuk mengkaji profil konsumen pabrik kulkas/ lemari es. Lembaga itu ingin mengetahui apakah ada relasi antara ukuran lemari es yang dibeli dengan jumlah anggota keluarga dari pembeli kulkas tersebut. Dari responden sejumlah 512 orang di kota besar diperoleh informasi sebagai berikut :

Ukuran lemari es yang dibeli	Banyak/Besar keluarga				Jumlah
	1-2	3-4	5-6	≥ 7	
Besar (12 lb load)	23	116	78	43	260
Medium (10 lb load)	54	25	16	11	106
Kecil (8 lb load)	31	68	39	8	146
Jumlah	108	209	133	62	512

Frekuensi yang diharapkan atau frekuensi teoretis (E_{ij}) dihitung sebagai

berikut:

$$\begin{aligned}
 E_{11} &= (260)(108)/512 = 54,6 \\
 E_{12} &= (260)(209)/512 = 106,1 \\
 E_{13} &= (260)(133)/512 = 67,5 \\
 E_{14} &= (260)(62)/512 = 31,5 \\
 E_{21} &= (160)(108)/512 = 22,4 \\
 E_{22} &= (160)(209)/512 = 43,3 \\
 E_{23} &= (160)(133)/512 = 27,5 \\
 E_{24} &= (160)(62)/512 = 12,8 \\
 E_{31} &= (146)(108)/512 = 30,8 \\
 E_{32} &= (146)(209)/512 = 59,6 \\
 E_{33} &= (146)(133)/512 = 37,9 \\
 E_{34} &= (146)(62)/512 = 17,7
 \end{aligned}$$

Nilai chi-square (X^2) dihitung sebagai berikut:

Frekuensi Hasil observasi dan Frekuensi yang Diharapkan dari Hubungan Besar Keluarga dengan Ukuran-ukuran lemari es

Ukuran lemari es yang Dibeli	Banyak/Besar keluarga				Jumlah
	1-2	3-4	5-6	≥ 7	
Besar (12 lb load)	23	116	78	43	260
Medium (10 lb load)	54	25	16	11	106
Kecil (8 lb load)	31	68	39	8	146
Jumlah	108	209	133	62	512

Keterangan:

Angka dalam kurung menunjukkan frekuensi yang diharapkan (E_{ij})

$$\begin{aligned}
 X^2 &= \sum_{i=1}^r \sum_{j=1}^k \frac{(O_{ij}-E_{ij})^2}{E_{ij}} \\
 &= \frac{(23-54,8)^2}{54,8} + \frac{(116-106,1)^2}{106,1} + \frac{(78-67,5)^2}{67,5} + \frac{(43-31,5)^2}{31,5} + \frac{(54-22,4)^2}{22,4} + \frac{(25-43,3)^2}{43,3} + \\
 &\frac{(16-27,5)^2}{27,5} + \frac{(11-12,8)^2}{12,8} + \frac{(31-30,8)^2}{30,8} + \frac{(68-59,6)^2}{59,6} + \frac{(39-37,9)^2}{37,9} + \frac{(8-17,7)^2}{17,7} \\
 &= 88,919
 \end{aligned}$$

$$C = \sqrt{\frac{X^2}{X^2 + n}}$$

$$C = \sqrt{\frac{88,919}{88,919 + 512}} = 0,385$$

Dengan demikian, koefisien yang dinyatakan dengan koefisien kontingensi antara besar keluarga dengan ukuran lemari es yang dibeli adalah $C = 0,385$.

Andai suatu kelompok subjek merupakan suatu sampel acak dari populasi tertentu, kita dapat menentukan apakah hubungan antara dua himpunan skor dari sampel itu menggambarkan bahwa memang nyata ada hubungan dalam populasinya, yaitu dengan jalan menguji signifikansi dari keterkaitan itu. Dalam menguji signifikansi korelasi tersebut, kita hendak menguji hipotesis nihil bahwa tidak terdapat korelasi dalam populasinya. Harga observasi ukuran korelasi dalam sampel hanya terjadi secara kebetulan dalam sebuah sampel random yang ditarik dari suatu populasi di mana kedua himpunan skor itu tidak berkorelasi.

Besarnya ukuran tabel kontingensi $r \times k$, kita dapat menentukan signifikansi koefisien kontingensi C dengan menetapkan probabilitas yang berkaitan dengan terjadinya, di bawah H_n , harga-harga sebesar harga observasi x^2 dengan $db = (r - 1)(k - 1)$. Apabila probabilitas itu lebih kecil dari (taraf signifikansi yang digunakan dalam pengujian), maka hipotesis nihil ditolak pada taraf signifikansi tersebut. Lampiran Tabel C (Tabel Harga *Chi-Square*) menyajikan probabilitas yang berkaitan dengan terjadinya, di bawah H_0 , harga-harga sebesar x^2 yang observasi. Jika x^2 untuk harga-harga sampel itu signifikan, maka kita dapat menyimpulkan bahwa dalam populasinya korelasi antara kedua himpunan skor tidak sama dengan nol.

Dari contoh di muka, yakni hubungan antara besar keluarga dengan ukuran lemari es yang dibeli, kita dapat menguji signifikansi $X^2 = 88,919$ dengan $db = (3 - 1)(4 - 1) = 6$ mempunyai probabilitas kemunculan di bawah H_0 yang besarnya.

Kurang (lebih kecil) dari 0,001. Dengan demikian, dapat kita putuskan bahwa H_0 ditolak pada taraf signifikansi (α) 0,01 dan menyimpulkan, bahwa terdapat hubungan antara besar keluarga dengan ukuran lemari es yang dibeli

dalam populasinya. Atau kita dapat menyimpulkan bahwa $C = 0,385$ secara signifikan berbeda dengan nol.

9.14 Korelasi Rank Spearman

Dalam penelitian, model korelasi rank (jenjang) Spearman digunakan untuk mengukur relasi antara dua variabel di mana dua variabel itu tidak mempunyai *joint normal distribution* dan *conditional varian* nya tidak sama. Korelasi rank berguna bila pengukuran kuantitatif secara eksak mustahil atau sulit dilakukan untuk mengukur koefisien korelasinya, maka disyaratkan bahwa pengukuran kedua variabelnya minimal dalam skala ordinal sehingga individu yang diamati dapat diberi jenjang dalam dua rangkaian berurutan.

Dalam analisis ini, hipotesis nihil yang akan diuji mengatakan bahwa dua variabel yang diteliti dengan nilai jenjangnya itu independen, artinya bahwa tidak ada hubungan antara jenjang variabel yang satu dengan jenjang yang lainnya. Pengujian dapat didasarkan pada sampel kecil ataupun sampel besar (apabila $n \geq 10$). Untuk sampel kecil digunakan tabel 0 pada lampiran, sedangkan untuk sampel besar dapat digunakan Lampiran Tabel B (Tab Harga-Harga Kritis t).

Koefisien korelasi rankspearman (dinotasikan dengan r) dinyatakan dengan rumus:

$$r_n = 1 - \frac{6 \sum_{j=1}^n d_j^2}{n(n^2-1)}$$

Di mana:

d = menunjukkan perbedaan setiap pasang rank.

n = menunjukkan jumlah pasangan rank.

Ukuran besarnya berbagai d tersebut akan menunjukkan seberapa erat hubungan antara skor kedua variabel. Jika hubungan antara kedua himpunan rank itu sempurna, setiap d , akan sama dengan nol.

Langkah-langkah dalam menggunakan analisis korelasi rank spearman dapat diikhtisarkan sebagai berikut:

1. Nilai pengamatan dari dua variabel yang akan diukur korelasinya diberi jenjang. Bila ada nilai pengamatan yang sama dihitung jenjang rata-ratanya!
2. Setiap pasang jenjang dihitung perbedaannya (d_i)!

3. Perbedaan setiap pasang jenjang tersebut dikuadratkan (d_1^2) dan dihitung jumlahnya!

4. Harga koefisien korelasi rankspearman dihitung dengan rumus:

$$r_n = 1 - \frac{6 \sum_{j=1}^n d_1^2}{n(n^2-1)}$$

5. Uji signifikansi harga observasi r, bergantung pada besar sampel:

a. Untuk sampel kecil gunakan Lampiran Tabel O!

b. Untuk sampel besar ($n \geq 10$), gunakan Lampiran Tabel B, di mana harga t observasi dihitung dengan rumus:

$$T = \frac{n-2}{1-r_n^2}$$

Ilustrasi Sampel Kecil:

Seorang dosen pendidikan olah raga yang menjadi pelatih cabang olah raga basket ingin melihat performa para atlet mahasiswa dengan memberikan peringkat performa saat simulasi/latihan dengan pertandingan yang sesungguhnya, pada rentang skala 0 hingga 100. Data yang diperoleh dari sampel beberapa pemain yang mengikuti pertandingan adalah sebagai berikut:

Pemain	Ranting	
	Dalam Latihan	Dalam Pertandingan
A	80	80
B	20	10
C	100	90
E	65	50
D	50	35
F	40	10
G	90	95
H	60	35

Prosedur pengujiannya adalah:

1. Formulasi H_0 dan H_1

$H_0: \rho = 0$ (tidak ada hubungan)

$H_1: \rho > 0$ (ada hubungan)

2. Digunakan uji signifikansi koefisien korelasi rank spearman.

3. Taraf signifikansi 0,01 dengan sampel $n = 8$. Lihat Lampiran Tabel O. Harga kritis r , untuk $n = 8$ dan $\alpha 0,01$ adalah 0,833.
4. Kriteria pengujian:
 H_0 ditolak apabila nilai rhoobservasi $> 0,833$ dan H_0 diterima apabila rhoobservasi $< 0,833$.
5. Nilai rho observasi (harga uji statistik r):

Pemain	Ranting		Rangking		d_1	d_1^2
	Dalam Latihan	Dalam Pertandingan	Dalam Latihan	Dalam Pertandingan		
A	80	80	6	6	0	0
B	20	10	1	1	0	0
C	100	90	8	7	1	1
E	65	50	5	5	0	0
D	50	35	3	3,5	0,5	0,25
F	40	10	2	2	0	0
G	90	95	7	8	-1	0
H	60	35	4	4	0,5	0,25
						2,5

$$r = 1 - \frac{6 \sum_{i=1}^n d_1^2}{n(n^2-1)}$$

$$r = 1 - \frac{6(6,25)}{8(8^2-1)} = 1-0,03 = 0,97$$

6. Keputusan pengujian:

Karena nilai rho observasi (0,97) lebih besar dari harga nilai kritis r (0,833) maka diputuskan H_0 ditolak (berarti H_1 diterima) pada taraf signifikansi 0,01. Dapat disimpulkan, bahwa terdapat derajat hubungan yang tinggi antara jenjang yang diberikan terhadap penampilan para pemain selama periode latihan dan jenjang selama pertandingan sesungguhnya.

Ilustrasi untuk Sampel Besar:

Dalam industri perakitan mobil, terutama merk – merk ternama dibutuhkan standar kecakapan tertentu. Dari kajian didapatkan data yang menunjukkan pengalaman kerja dalam merakit mobil (dalam minggu) dan jumlah rakitan mobil yang tidak memenuhi standar dalam unit dari sampel acak sejumlah 12 pekerja.

Sampel Pekerja	Pengalaman(Minggu)	Di Bawah Standar(unit)
1	7	26
2	9	20
3	6	28
4	14	10
5	8	23
6	12	18
7	10	24
8	4	26
9	2	38
10	11	22
11	1	32
12	8	25

Adakah hubungan antara pengalaman dengan produk yang tidak memenuhi standar?

Prosedur pengujiannya adalah:

1. Formulasi H_0 dan H_1
 $H_0: \rho = 0$ (tidak ada hubungan)
 $H_1: \rho < 0$ (ada hubungan negatif)
2. Digunakan uji signifikansi korelasi rankspearman.
3. Digunakan taraf signifikansi 0,05 dengan besar Sampel, $n = 12$. Lihat Lampiran Tabel B! Harga/nilai kritis $t = -1,812$ (satu sisi Kiri).
4. Kriteria pengujian:
 H_0 ditolak apabila nilai t observasi $< -1,812$ dan H_0 diterima apabila nilai t observasi $> -1,812$.
5. Harga/nilai t observasi:

Sampel Pekerja	Pengalaman(x)	Produk Rusak(Y)	Rank X	Rank Y	D_i	D_i^2
1	7	26	5	8,5	-3,5	12,25
2	9	20	8	3	5	25
3	6	28	4	10	6	36
4	14	10	12	1	11	121
5	8	23	6,5	5	1,5	2,25
6	12	18	11	2	9	81
7	10	24	9	6	3	9
8	4	26	3	8,5	-5,5	30,25
9	2	38	2	12	-10	100

10	11	22	10	4	6	36
11	1	32	1	11	-10	100
12	8	25	6,5	7	-0,5	0,25
						533

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)}$$

$$r = 1 - \frac{6(553)}{12(12^2-1)} = 1 - 1,93 = -0,93$$

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

$$t = -0,93 \sqrt{\frac{12-2}{1-(-0,93)^2}} = -8,0012$$

5. Keputusan pengujian:

Karena harga t observasi (-8,0012) lebih kecil dari harga kritis $t = -1,812$ maka H_0 ditolak (berarti H_1) diterima pada taraf signifikansi 0,5. Dapat disimpulkan, bahwa terdapat hubungan negatif antara pengalaman dengan jumlah produk yang tidak memenuhi standar, artinya semakin lama pengalaman kerja akan semakin sedikit jumlah produk yang tidak memenuhi standar/rusak.

SOAL – SOAL LATIHAN

1. Apa perbedaan elementer antara uji statistik parametrik dan nonparametrik? Jelaskan.
2. Ada beberapa asumsi yang bergayut kuat dengan uji statistic nonparametrik. Jelaskan dengan singkat!
3. Apa perbedaan mendasar antara korelasi rank Kendall dan spearman?
4. Agen tunggal pemegang merk sebuah kendaraan roda 4 ingin mengetahui jenis mobil yang digemari masyarakat. Melalui divisi riset pasar, dilakuan survei selama 30 hari terhadap konsumen dan didapatkan data sebagai berikut: mobil merk Toyota 1000 orang, Daihatzu 850 orang, Honda 650 Orang dan merk lainnya 350 Orang.
Merk apa yang paling di sukai orang?
5. Berapa prosentase mobil Honda disukai orang?

BAB X

KOMPUTASI DATA DENGAN PROGRAM SPSS

10.1 Sekilas Program SPSS

Pada zaman digital seperti saat ini, pekerjaan manusia menjadi mudah karena komputasi data yang telah berkembang pesat. Sekadar napak tilas, seiring dengan meluasnya penggunaan komputer, maka program untuk perhitungan statistik juga kian canggih. Dahulu kita hanya mengenal program berbasis DOS seperti Mirostat hingga kini kita sampai pada program yang berbasis Windows, seperti SPSS (*Statistical Product and Service Solutions*), Minitab, SAS (*Statistical Analysis System*), Statistica, Statistrx, S-Plus, dan sebagainya.

Pada masing – masing program tersebut, tentu mempunyai berbagai kelebihan dan kekurangan, yang sebenarnya sudah pula ditentukan kegunaannya. Kebutuhan perhitungan statistik perlu diperhatikan sehingga pengguna dalam hal ini peneliti tidak perlu untuk menuntut diri menguasai berbagai program tersebut. Dengan demikian dapat disarankan dan sebenarnya terserah juga pada para peneliti, bahwa cukup dengan menguasai satu program statistik tertentu saja, maka semua kebutuhan penghitungan statistik dapat dilakukan.

Bab ini hanya dengan singkat akan membicarakan pemanfaatan paket program SPSS, sebab program ini cukup bahkan sangat populer digunakan dalam berbagai disiplin ilmu untuk menyelesaikan aneka permasalahan dalam penelitian pasar, kajian ilmu – ilmu ekonomi, sosial, marketing, ilmu – ilmu eksata dan lainnya. Inilah yang pada akhirnya program SPSS menjadi multidisiplin dalam penggunaannya. SPSS pada mulanya adalah kependekan dari *Statistical Package for the Social Sciences* sekarang bertransformasi menjadi semakin melebar, yakni menjadi *Statistical Product and Service Solutions*.

Berbarengan dengan perkembangan yang sangat pesat dan kebutuhan yang begitu beragam, maka mulai semenjak 1998 program SPSS dikembangkan untuk memenuhi kebutuhan berdasar kategori yang disusun oleh pembuatnya menjadi empat macam sebagai berikut :

1. SPSS BI atau *Business Intelgense* untuk pasar bisnis;

2. SPSS MR atau Market Research untuk riset pasar;
3. SPSS Science untuk riset sains; dan
4. SPSS *Quality* untuk peningkatan kualitas.

10.2 Mengoperasikan Program SPSS

Untuk memulai operasi program SPSS *for windows* dapat mengikuti tahapan langkah sebagaimana berikut ini:

- 1) Harus dipastikan bahwa pada komputer yang akan digunakan sudah diinstal program paket SPSS.
- 2) Lalu klik **START**.
- 3) Arahkan *pointer mouse* ke Program, kemudian rahkan kembali *pointer mouse* ke **SPSS for Window**.
- 4) Lalu klik *icon* IBM **SPSS Statistics**.

Setelah langkah-langkah di atas dilakukan maka akan berhadapan dengan layar SPSS, yaitu **Data Editor**. Pada tampilan pertama muncul dialog. Jika kita sudah memiliki data yang siap untuk dioperasikan dengan SPSS, maka pilih data yang akan diolah pada kotak dialog, kemudian **klik OK**. Tetapi jika belum ada data yang tersedia pada kotak dialog tersebut maka **klik Cancel**.

Untuk mengakhiri SPSS dapat dilakukan dengan **klik File** pada baris menu, kemudian **klik Exit**, atau **klik icon X** pada sudut kanan atas.

10.3 Mengenal Windows SPSS

Pada langkah awal, yang dapat dilakukan adalah mengenali aplikasi berbasis windows yang didalamnya telah diinstal SPSS. Didalam paket program itu terdapat 4 window, yaitu Data Editor, Output Viewer, Syntax Editor, dan Script Editor.

Menu-menu yang tersedia pada Data Editor adalah:

a. *Menu File*

Menu **file** ini berfungsi untuk menangani masalah yang berhubungan dengan file data seperti:

New (Data, Syntax, Output, Draft Output, Script) membuat file baru

Open (Data, Syntax, Output, Draft Output, Script)membuka file baru

Open Database (New Query, Edit Query, Run Query) mengimpor data dari Program lain, seperti Excel, SAS, Minitab, dan sebagainya.

Read Text Data membaca data text

Save menyimpan data yang sedang aktif ke dalam file SPSS

Save As menyimpan data yang sedang aktif ke dalam berbagai tipe data.

Display Data Info membuka beberapa nama file yang sudah ada dalam bentuk SPSS.

Print mencetak data yang sedang aktif

Print Preview melihat data yang sedang aktif sebelum dicetak.

Stop Processor menghentikan/membatalkan perintah yang sedang berjalan.

b. Menu Edit

Menu **Edit** ini berfungsi untuk menangani hal-hal yang berhubungan dengan memperbaiki atau mengubah nilai data, seperti:

Undo memanggil ulang perintah yang sudah berlalu.

Redo memanggil kembali perintah yang sudah dibuat.

Cut menghapus text atau data yang disorot.

Copy menggandakan text atau data yang disorot ke *Clipboard*

Paste menempatkan text atau data yang dicopy ke Clipboard.

Clear menghapus text atau data yang disorot.

Find mencari data atau text yang diinginkan.

Options mengubah setting, seperti *output* tabel, label, *script*, dan sebagainya,

c. Menu View

Menu **View** ini berfungsi untuk mengatur toolbar, seperti:

Status Bar mengaktifkan status bar.

Toolbars mengatur penampilan toolbar yang ada pada SPSS.

Font mengatur ukuran dan bentuk huruf yang akan digunakan.

Grid Lines mengaktifkan garis garis pada sel data

Value Label mengaktifkan/menampakkan nilai pada label.

d. Menu Data

Menu **Data** ini berfungsi untuk membuat perubahan data SPSS secara keseluruhan, seperti:

Define Dates mengerjakan pengisian variabel yang berhubungan dengan deret waktu.

Insert Variable menyisipkan variabel.

Insert Cases menyisipkan kasus pada sel tertentu.

Go to Cases menemukan sel tertentu.

Short Cases mengurutkan data yang disorot.

Transposed mentransformasi data dari baris menjadi kolom atau sebaliknya.

Merger File menggabungkan file yang sekarang ada dengan file baru.

Aggregate meringkas data secara agregat.

Split File memisahkan isi file dengan kriteria tertentu.

Select Cases menyeleksi file dengan kriteria tertentu.

Weight Cases memberi bobot pada suatu kasus tertentu.

e. Menu Transform

Menu **Transform** ini berfungsi untuk membuat perubahan pada variabel yang telah dipilih dengan kriteria tertentu, seperti:

Compute menambah variabel baru yang berisi hasil perhitungan berdasarkan data dari variabel lama.

Random Number Seed digunakan untuk menciptakan sejumlah angka acak untuk keperluan tertentu.

Count menghitung data dengan kriteria tertentu.

Recode memberi kode ulang ke suatu variabel berdasarkan kriteria tertentu, baik dalam variabel yang sama maupun dalam variabel yang berbeda.

Categorize Variables mengubah bilangan numerik yang kontinu menjadi bilangan diskrit atau kategorik.

Rank Cases mengurutkan kasus dengan kriteria tertentu.

Automatic Recode mengubah sebuah variabel numerik tanpa nilai menjadi sebuah variabel numerik dengan nilai secara otomatis.

Create Time Series fungsi ini berhubungan dengan pembuatan variabel pada data Time Series.

Replace Missing Value fungsi ini berhubungan dengan pengelolaan nilai *missing* pada data *Time Series*.

f. Menu Analyze

Menu **Analyze** ini berfungsi untuk melakukan semua prosedur perhitungan statistika. Menu ini adalah inti dari modul ini sehingga akan dibahas dalam bab tersendiri.

g. Menu Graphs

Menu **Graphs** ini berfungsi untuk membuat berbagai jenis grafik yang dikehendaki, seperti:

Status Bar mengaktifkan status bar.

h. Menu Utilities

Menu **Utilities** atau menu tambahan yang mendukung program SPSS, seperti: **Variables** member informasi tentang sebuah variabel yang disorot.

File Info memberi informasi tentang semua variabel yang ada pada data yang sedang aktif.

Define Set mengubah atau menampilkan variabel- variabel yang dianggap perlu saja.

Use Set mengubah atau menampilkan variabel-variabel yang dianggap perlu saja.

Menu Editor menyediakan fasilitas untuk dapat menampilkan menu tambahan tersendiri di luar menu-menu standar yang sudah ada.

i. Menu Window

Menu **Window** ini berfungsi untuk berpindah di antara menu-menu lain yang ada dalam SPSS.

j. Menu Help

Menu **Help** ini berfungsi untuk menyediakan bantuan informasi mengenai program SPSS yang dapat diakses secara mudah dan jelas. Menu ini berisi antara lain:

➤ **TOPICS** untuk melihat tiap topik mengenai cara kerja SPSS.

- **TUTORIAL** untuk melihat tiap topik mengenai cara kerja SPSS, dapat tiap topik yang khusus dengan cara mencari kata kunci lewat INDEX.
- **STATISTICS COACH** untuk melihat tiap topik statistik yang diperlukan dan kaitannya dengan pengerjaan SPSS.

10.4 Input (Memasukkan) Data

Poses dalam input data pada program SPSS dapat dilakukan dengan berbagai cara, yaitu:

- a. Memanggil data dari program lain, jika data sudah tersimpan dalam program lain, seperti (Excel, Minitab Systat, dBase, Lotus, dan lain lain).
- b. Memanggil data yang sudah pernah dibuat atau sudah ada dalam program SPSS. Langkahnya bisa ditelusuri hanya file datanya sudah tersimpan dalam program SPSS (*.sav)
- c. Memasukkan data langsung dalam window (sel) SPSS pada Data Editor. Langkah-langkahnya adalah:

1. Buka lembar kerja baru (klik menu utama **File**,sorot menu **New**, klik sub menu **Data**)
2. Persiapkan nama dan karakteristik variabel dengan menampilkan VARIABEL VIEW (klik tab sheet **variabel View** yang ada di bagian kiri bawah atau tekan **CTRL + T**)
3. Setelah terlihat tampilan SPSS DATA EDITOR,kemudian berinama variabel yang diperlukan dan karakteristik variabel tersebut.

10.5 Analisis Faktor dengan SPSS for Window

Langkah-langkah analisis faktor dengan SPSS adalah sebagai berikut:

1. Pastikann bahwa program SPSS sudah ada dalam komputer (semua versi SPSS dapat digunakan).
2. Klik **Start>All Program>IBM SPSS Statistics**
3. Setelah Windows SPSS terbuka dan aktif, masukkan data pada '**Data View**' dan beri penjelasan masing-masing variabel berupa nama variabel, label, dan sebagainya dengan mengaktifkan Windows '**Variable View**' Teknik penulisan datanya dilakukan dengan menuliskan ke bawah untuk responden dan ke samping untuk semua item/variabel.

4. Setelah data siap artinya semua responden dan semua item dari kuesioner telah masuk semua dan telah diberi label pada masing masing item/variabel, selanjutnya mulai dengan analisis faktor dengan langkah; klik **Analyze > Data Reduction > Factor ...** sehingga muncul kotak dialog.
5. Masukkan semua item/variabel yang akan dianalisis ke dalam kotak **Variable** dengan cara: blok semua item/variabel (misal: x1 - x30), kemudian klik tanda ► sehingga variabel tadi (x1 - x30) masuk dalam kotak Variable. Selanjutnya klik **Descriptives...** sehingga muncul kotak dialog. Tandai nilai-nilai yang ingin muncul dalam output, kemudian klik **Continue**.
6. Kemudian klik **Extraction..** sehingga muncul kotak dialog.. Pilih metode yang akan digunakan dengan cara klik tanda ↗ pada kotak **Method**, misal *Principal components*, kemudian tandai **Number of factors** jika kita mengharapkan ke-30 variabel tadi ingin dibentuk menjadi 5 faktor, atau biarkan saya maka SPSS akan membentuk jumlah faktor sebanyak faktor yang memihki nilai *eigen* lebih dar 1 (*Default*). Kemudian klik **Continue**,
7. Berikutnya klik **Rotation**, sehingga muncul Kotak dialog.. Pilih metode rotasi yang akan digunakan, misal **Varimax**, maka tandai pada metode Varimax. Kemudian pada **Display** biarkan seperti **Default**, tapi jika mengharapkan plot loadingnya, maka tandai **Loading plot (s)**. Kemudian klik **Continue**.
8. Berikutnya klik **Scores . . .**, sehingga muncul kotak dialog seperti pada jika kita mengharapkan nilai skor faktornya disimpan dalam data view, maka tandai **Save as variables**, kemudian pilih metode yang digunakan dalam menduga skor faktor tersebut. Misalkan pilih metode pendugaan dengan regresi, maka tandai **Regression** pada kotak tersebut, kemudian klik **Continue**.
9. Berikutnya klik **Options . . .**, sehingga muncul kotak dialog. Pada kotak dialog ini terdiri dari dua bagtan, yaitu tentang nilai yang hilang (**Messsing values**) dan format koefisien yang akan ditampilkan. Pada missing values biarkan seperti **Default**, sedangkan untuk format koefisien yang ingin ditampilkan keduanya diberi tanda, yaitu untuk **Sorted by size**, artinya koefisien tersebut akan

ditampilkan berurutan dari yang memiliki faktor *loading* terbesar sampai terkecil, dan **Suppress absolute values less than** kemudian isi kan koefisien minimal yang diharapkan pada kotak misal 0.50, artinya akan membuang semua koefisien yang lebih kecil dari 0,50. jika kedua bagian ini tidak diberi tanda, maka *output* hasil akan memberikan tampilan urutan variabel berdasarkan nama variabel, dan semua koefisien yang ada pada masing-masing faktor akan dikeluarkan. Kemudian klik **Continue**.

10. Kemudian klik **OK** sehingga muncul **windows output**. Sedangkan nilai-nilai dari **skor faktor** akan tersimpan dalam windows data pada **data view**.
11. Kemudian pada Variabel view beri nama pada masing-masing faktor sesuai dengan item/variabel yang masuk dalam faktor tersebut

10.6 Analisis Korelasi dengan SPSS for Windows

1. Analisis Korelasi Sederhana dengan SPSS for Windows

Langkah-langkah dalam analisis korelasi sederhana untuk melihat hubungan antara tingkat motivasi dengan kepuasan kerja adalah sebagai berikut:

1. Buka lembar kerja dan pastikan data sudah tersedia dengan kriteria yang sudah ditentukan.
2. Dari menu utama SPSS pilih menu **Analyze**, kemudian pilih submenu **Correlate**.
3. Sorot ke kanan **Bivariate...**, kemudian klik pilihan tersebut sehingga muncul kotak dialog
4. Kemudian pilih variabel **motivasi** dan **kepuasan** sehingga masuk ke kotak **Variables**.
5. Pada bagian **Correlation Coefficients** pilih **Pearson**, karena datanya dalam skala interval. Jika datanya ordinal maka pilihannya adalah **Spearman**.
6. Bagian **Test of Significance** pilih **Two tailed** (dua arah) karena hipotesisnya menggunakan uji dua arah.
7. Kemudian klik **OK**, sehingga diperoleh *output* pada **window output**.
8. Hasil *output* dari analisis korelasi **pearson** adalah.

		Motivasi	Kepuasan
Motivasi	Pearson Correlation	1000	695**
	Sig.(2-tailed)	-	.000
	N	43	43
Kepuasan	Pearson Correlation	695**	1000
	Sig.(2-tailed)	.000	-
	N	43	43

** Correlation is significant at the 0.01 level (2-tailed)

9. Analisis dan Interpretasi.

Output di atas menunjukkan nilai koefisien korelasi sebesar 0,695. Hasil pengujian menunjukkan p -value (Sig) sebesar 0,0001. Karena p -value $< 5\%$, maka H_0 ditolak. Artinya, dengan taraf nyata sebesar 5% dapat dinyatakan bahwa ada hubungan yang nyata antara motivasi kerja dengan kepuasan kerja pegawai, meskipun hubungannya moderat.

2. Analisis Korelasi Parsial dengan SPSS for Windows

Dengan langkah yang sama seperti analisis korelasi sederhana pada analisis korelasi parsial diperoleh hasil analisis dan interpretasi sebagai berikut:

- Hipotesis yang akan diuji untuk melihat hubungan motivasi kerja (X_1) dengan kepuasan kerja (Y) jika variabel disiplin kerja (X_2) tetap adalah:
 - $H_0 : \rho_{yx_1(x_2)} = 0$ artinya tidak ada hubungan antara variabel kepuasan kerja dengan variabel motivasi kerja, jika variabel disiplin kerja dianggap tetap.
 - $H_1 : \rho_{yx_1(x_2)} \neq 0$ artinya ada hubungan antara variabel kepuasan kerja dengan variabel motivasi kerja, jika variabel disiplin kerja dianggap tetap.
- Statistik uji yang akan digunakan adalah uji-t, dengan kriteria uji tolak H_0 jika p -value $< \alpha$.
- Untuk menguji hipotesis tersebut, akan digunakan analisis korelasi parsial dengan paket Program SPSS versi 14.0.

4. Dari hasil di atas dapat dijelaskan sebagai berikut:

Output di atas menunjukkan nilai korelasi parsial sebesar 0,5741. Karena $p\text{-value} = 0.000 < 5\%$, maka H_0 ditolak. Artinya, ada hubungan positif antara variabel kepuasan kerja dengan motivasi kerja sebesar 57,41%, jika variabel disiplin dianggap tetap.

10.7 Analisis Regresi dengan SPSS for Windows

Langkah langkah dalam analisis regresi berganda untuk melihat pengaruh kepemimpinan dan disiplin kerja terhadap kinerja pegawai adalah sebagai berikut:

1. Buka lembar kerja dan pastikan data sudah tersedia dengan kriteria yang sudah ditentukan.
2. Dari menu utama SPSS, pilih menu **Analyze**, kemudian pilih submenu **Regression**.
3. Sorot ke kanan **Linear**, kemudian klik pilihan tersebut sehingga muncul kotak dialog
4. Kemudian pilih variabel *kinerja pegawai* sebagai variabel terikat sehingga masuk ke kotak **Dependent**.
5. Kemudian pilih variabel bebas *kepemimpinan* dan *disiplin kerja* sehingga masuk ke kotak **Independents**.
6. Pada bagian lainnya diisi sesuai dengan kebutuhan yang ingin diungkapkan atau dikeluarkan.
7. Hasil *output* nya adalah sebagai berikut:

OUTPUT SPSS

Regression

Descriptive Statistics

	Mean	Std.Deviation	N
Kinerja	2.546883	.487384	36
Pegawai	2.898573	.455657	36
Kepimpinan	2.755381	.481213	36
Disiplin Kerja			

Variable Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	Disiplin Kerja, Kepemimpinan		Enter

- All requested variables entered
- Dependent variable: Kinerja Pegawai

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.760 ^a	.578	.552	.326134

- Predictors: (Constant), Disiplin kerja, Kepemimpinan
- Dependent Variable: Kerja Pegawai

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.804	2	2.402	22.583	.000 ^a
	Residual	3.510	33	.106		
	Total	8.314	35			

- Predictors: (Constant), Disiplin kerja, Kepemimpinan
- Dependent Variable: Kerja Pegawai

10.8 Penyajian Hasil Analisis dan Interpretasi

Salah satu hal yang paling penting dalam proses analisis data adalah cara penyajian dan interpretasi dari hasil pengolahan data. *Output-output* yang dikeluarkan dari program SPSS hanyalah berupa tabel dan angka, tanpa ada penjelasan atau interpretasi. Dengan demikian, *output* yang diperoleh dari SPSS hendaknya dipindahkan ke dalam program pengolah kata, seperti MS Word, MS Power Point, atau yang lain untuk memberikan penjelasan dari hasil tersebut. Langkah-langkah memindahkan *output* dari Windows SPSS ke Windows MS Word adalah :

- Buka **Windows Output** SPSS.
- Klik (blok bagian **output** SPSS yang akan d copy).
- Buka menu utama **Edit**, kemudian klik **Copy objects**.

Atau dengan cara *klik* kanan pada mouse, kemudian *klik* kiri bagian *copy objects*, seperti pada

Atau setelah diblok tekan **Ctrl + K**

1. Buka windows MS Word, tempatkan kursor yang akan dijadikan tempat output SPSS. Buka menu utama Edit, kemudian klik Paste. Atau setelah kursor ditempatkan kemudian tekan Ctrl + V sehingga muncul output hasil SPSS pada Windows MS Word.
2. Buatlah format atau sesuaikan dengan tampilan yang dikehendaki serta beri penjelasan dari output tersebut.

Contoh hasil analisis regresi berganda dengan program SPSS di atas,

Tabel Nilai Koefisien Determinasi Analisis Regresi Berganda Model

Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.760 ^a	.578	.552	.326134

Predictors: (Constant), Disiplin Kerja, Kepimpinan

Pada tabel di atas (Model Summary) diperoleh nilai $R = 0,760$. Artinya, ada hubungan korelasi positif antara variabel kepemimpinan dan disiplin kerja dengan tingkat kinerja pegawai sebesar 76%. Nilai koefisien determinasi atau R^2 (R Square) = 0,578 menunjukkan bahwa besarnya kontribusi pengaruh variabel kepemimpinan (X_1) dan disiplin kerja (X_2) terhadap kinerja pegawai (Y) sebesar 57,8%

Tabel Analisis Varian (Analysis of Varians)

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.804	2	2.402	22.58	.000 ^a
	Residual	3.510	33	.106	3	
	Total	8.314	35			

a. Predictors: (Constant), Disiplin kerja, Kepimpinan

b. Dependent Variable: Kerja Pegawai

Pada Tabel (ANOVA) menunjukkan pengujian secara simultan untuk model regresi yang melibatkan variabel X_1 dan X_2 terhadap Y . Dari hasil pengujian diperoleh nilai F -hitung sebesar 22,583 dengan p -value (sig) = 0,0001. Karena p -value (sig.) < 5%, maka H_0 ditolak, Artinya, dengan tingkat kesalahan sebesar 5% dapat dinyatakan bahwa kepemimpinan dan disiplin kerja memiliki pengaruh nyata terhadap tingkat kinerja pegawai.

Tabel Nilai Koefisien Regresi Berganda

Coefficients^a

Model	Unstandardized		t	Sig
	Coefficients			
	B	Std. Error		
1 (Constant)	.105	.378	.278	.782
Kepemimpinan	.324	.151	2.139	.040
Disiplin Kerja	.545	.143	3.802	.001

a. Dependent Variable: Kinerja Pegawai

Tabel 11.3 pada kolom (Coefficients) menunjukkan pengujian secara parsial dengan uji-t.

- a. Hasil pengujian untuk variabel kepemimpinan diperoleh koefisien atau nilai $B = 0,324$, dan nilai t -hitung = 2 139 atau p -value (Sig) = 0,040. Karena p -value < 5%, maka H_0 ditolak Artinya, dengan tingkat kesalahan paling besar 5% dapat dinyatakan bahwa kepemimpinan berpengaruh nyata terhadap tingkat kinerja pegawai.
- b. Hasil pengujian untuk variabel disiplin kerja diperoleh koefisien atau nilai $B = 0,545$, dan nilai t -hitung = 3,802 atau p -value (Sig.) = 0,001. Karena p -value < 5%, maka H_0 ditolak. Artinya, dengan tingkat kesalahan paling besar 5% dapat nyatakan bahwa disiplin kerja berpengaruh nyata terhadap tingkat kinerja pegawai.

DAFTAR PUSTAKA

- Cooper, D.R dan C.W Emory (1991) *Statistic for Management*. Singapore :
Wadsworth Asian Student Edition
- Dajan, Anto (2000) *Pengantar Metode Statistik Jilid II*, Jakarta: LP3ES
- Deibler, F. S (1943) Horace Secrist, 1881–1943, *Journal of the American Statistical Association*, 38:223, 365366, DOI: [10.1080/01621459.1943.10501821](https://doi.org/10.1080/01621459.1943.10501821)
- Malo, M (1986) *Metode Penelitian Sosial*. Universitas Terbuka. Jakarta:
Karunika
- Siegel, S. (1956) *Non Parametric Methods for the Behavioral Sciences*. New
York: McGraw-Hill
- Siegel, S. (1982) *Non Parametric Methods for the Behavioral Sciences*. New
York: McGraw-Hill
- Suriasumantri, Jujun S. 2003. *Filsafat Ilmu Sebuah Pengantar Populer*.
Jakarta: Pustaka Sinar Harapan.
- Tabachnik, B,BG dan L.S Fidell (1996) *Using Multivariate Statistics*. New
York : Harper Collins College Publishers.
- UU No.20/2008 tentang Usaha Mikro, Kecil, dan Menengah
<https://asahankab.go.id/>

DAFTAR TABEL

Tabel III :

Tabel Harga Kritis U untuk Uji Satu Sisi pada $\alpha = 0,025$, atau untuk Uji Dua Sisi pada $\alpha = 0,05$.

n1\n	9	10	11	12	13	14	15	16	17	18	19	20
2												
1												
2	0	0	0	1	1	1	1	1	2	2	2	2
3	2	3	3	4	4	5	5	6	6	7	7	8
4	4	5	6	7	8	9	10	11	11	12	13	13
5	7	8	9	11	12	13	14	15	17	18	19	20
6	10	11	13	14	16	17	19	21	22	24	25	27
7	12	14	16	18	20	22	24	26	28	30	32	34
8	15	17	19	22	24	26	29	31	34	36	38	41
9	17	20	23	26	28	31	34	37	39	42	45	48
10	20	23	26	29	33	36	39	42	45	48	52	55
11	23	26	30	33	37	40	44	47	51	55	58	62
12	26	29	33	37	41	45	49	53	57	61	65	69
13	28	33	37	41	45	50	54	59	63	67	72	76
14	31	36	40	45	50	55	59	64	67	74	78	83
15	34	39	44	49	54	59	64	70	75	80	85	90
16	37	42	47	53	59	64	70	75	81	86	92	98
17	39	45	51	57	63	67	75	81	87	93	99	105
18	42	48	55	61	67	74	80	86	93	99	106	112
19	45	52	58	65	72	78	85	92	99	106	113	119
20	48	55	62	69	76	83	90	98	105	112	119	127

Tabel Jiv :
Tabel Harga Kritis U untuk Uji Satu Sisi pada $\alpha = 0,05$, atau untuk Uji Dua Sisi pada $\alpha = 0,10$

n1\n2	9	10	11	12	13	14	15	16	17	18	19	20
1											0	0
2	1	1	1	2	2	2	3	3	3	4	4	4
3	4	4	5	5	6	7	7	8	9	9	10	11
4	6	7	8	9	10	11	12	14	15	16	17	18
5	9	11	12	13	15	16	18	19	20	22	23	25
6	12	14	16	17	19	21	23	25	26	28	30	32
7	15	17	19	21	224	26	28	30	33	35	37	39
8	18	20	23	26	28	31	33	36	39	41	44	47
9	21	24	27	30	33	36	39	42	45	48	51	54
10	24	27	31	34	37	41	44	48	51	55	58	62
11	27	31	34	38	42	46	50	54	57	61	65	69
12	30	34	38	42	47	51	55	60	64	68	72	77
13	33	37	42	47	51	56	61	65	70	75	80	84
14	36	41	46	51	56	61	66	71	77	82	87	92
15	39	44	50	55	61	66	72	77	83	88	94	100
16	42	48	54	60	65	71	77	83	89	95	101	107
17	45	51	57	64	70	77	83	89	96	102	109	115
18	48	55	61	68	75	82	88	95	102	109	116	123
19	51	58	65	72	80	87	94	101	109	116	123	130
20	54	62	69	77	84	92	100	107	115	123	130	138

Tabel K :
Tabel Harga-Harga Kritis K_D dalam Uji Dua Sampel
Kolmogorov-Smirnov (Sampel-sampel Kecil)

n	One – tailed Test		Two- tailede Test	
	$\alpha = .05$	$\alpha =.01$	$\alpha = .05$	$\alpha =.01$
3	3			
4	4		4	
5	4	5	5	
6	5	6	5	5
7	5	6	6	6
8	5	6	6	6
9	6	7	6	7
10	6	7	7	7
11	6	8	7	8
12	7	8	7	8
13	7	8	8	8
14	7	8	8	9
15	7	9	8	9
16	8	9	8	9
17	8	9	8	10
18	8	10	9	10
19	8	10	9	10
20	8	10	9	10
21	8	10	9	11
22	9	11	9	11
23	9	11	10	11
24	9	11	10	11
25	9	11	10	12
26	9	11	10	12
27	9	12	10	12
28	10	12	11	12
29	10	12	11	13
30	10	12	11	13

35	11	13	12	13
40	11	14	13	

Tabel Harga-Harga Kritis D dalam uji Dua Sampel Kolmogorov-Smirnov(Sampel-Sampel Besar :Uji Dua Sisi

Level of significance	Value of so large as to call for rejection H_0 At the indicated level of significance where $D = \text{maximum } S_{n1}(X) - S_{n2}(X) $
.10	$1.22 \sqrt{\frac{n_1+n_2}{n_1 n_2}}$
.05	$1.36 \sqrt{\frac{n_1+n_2}{n_1 n_2}}$
.025	$1.48 \sqrt{\frac{n_1+n_2}{n_1 n_2}}$
.01	$1.63 \sqrt{\frac{n_1+n_2}{n_1 n_2}}$
.005	$1.73 \sqrt{\frac{n_1+n_2}{n_1 n_2}}$
.001	$1.95 \sqrt{\frac{n_1+n_2}{n_1 n_2}}$

Tabel L :
Tabel Harga-Harga Kritis D dalam Uji Dua Sampel Kolmogorov-
Smirnov(Sampel-Sample Besar : Uji Dua Sisi)

n=2		n=3		n=4		n=5	
X_r^2	P	X_r^2	P	X_r^2	P	X_r^2	P
0	1.000	.000	1.000	.0	1.000	.0	1.000
1	.833	.667	.944	.5	.931	.4	.954
4	.500	2.000	.528	1.5	.653	1.2	.691
3	.167	2.667	.361	2.0	.431	1.6	.522
		4.4667	.194	3.5	.273	2.8	.367
		6.000	.028	4.5	.125	3.6	.182
				6.0	.069	4.8	.124
				6.5	.042	5.2	.093
				8.0	.0046	6.4	.039
						7.6	.024
						8.4	.0085
						10.0	.00077

n=6		n=7		n=9		n=9	
X_r^2	P	X_r^2	P	X_r^2	P	X_r^2	P
.00	1.000	.000	1.000	.00	1.000	.000	1.000
1.33	.956	.286	.964	.25	.967	.222	.971
1.00	.740	.857	.768	.75	.794	.667	.814
2.33	.570	.143	.620	1.00	.654	.889	.865
2.33	.430	.000	.486	1.75	.531	1.556	.569
3.00	.252	.571	.305	2.25	.355	2.000	.398
4.00	.184	3.429	.237	3.00	.285	2.667	.328
4.33	.142	3.714	.192	3.25	.236	2.889	.278
5.33	.072	4.571	.112	4.00	.149	3.556	.187
6.33	.052	5.429	.085	4.75	.120	4.222	.154
7.00	.029	6.000	.052	5.25	.079	4.667	.107
8.33	.012	7.143	.027	6.25	.047	5.556	.069
9.00	.0081	7.714	.021	6.75	.038	6.000	.057

9.33	.0055	8.000	.016	7.00	.030	6.222	.048
10.33	.0017	8.857	.0084	7.75	.018	6.889	.031
12.00	.00013	10.286	.0036	9.00	.0099	8.000	.019
		10.571	.0027	9.25	.0080	8.222	.016
		11.143	.0012	9.75	.0048	8.667	.010
		12.286	.00032	10.75	.0024	9.556	.0060
		14.000	.000021	12.00	.0011	10.667	.0035
				12.25	.00086	10.889	.0029
				13.00	.00026	11.556	.0013
				14.25	.000061	12.667	.00066
				16.00	.0000036	13.556	.00035
						14.000	.00020
						14.222	.000097
						14.889	.000054
						16.222	.000011
						18.000	.000000

n=2		n=3		n=4			
X_r^2	P	X_r^2	P	X_r^2	P	X_r^2	P
.0	1.000	.2	1.000	.0	1.	5.7	.141
.6	.958	.6	.958	.3	.992	6.0	.105
1.2	.834	1.0	.910	.6	.928	6.3	.094
1.8	.792	1.8	.727	.9	.900	6.6	.077
2.4	.625	2.2	.608	1.2	.800	6.9	.068
3.0	.542	2.6	.524	1.5	.754	7.2	.054
3.6	.458	3.4	.446	1.8	.677	7.5	.052
4.2	.375	3.8	.342	2.1	.649	7.8	.036
4.8	.208	4.2	.300	2.4	.524	8.1	.033
5.6	.167	5.0	.207	2.7	.508	8.4	.019
6.0	.042	5.4	.175	3.0	.432	8.7	.014
		5.8	.148	3.3	.389	9.3	.012
		6.6	.075	3.6	.355	9.6	.0069
		7.0	.054	3.9	.324	9.9	.0062
		7.4	.033	4.5	.242	10.2	.0027

		8.2	.017	4.8	.200	10.8	.0016
		9.0	.0017	5.1	.190	11.1	.00094
				5.4	.158	12.0	.000072

Tabel N :
Tabel Kemungkinan yang Berkaitan dengan Harga-Harga sebesar Harga-Harga Observasi H dalam Analisis Varian Ranking Satu Arah Kruskal-Wallis

Ukuran Sampel			H	P	Ukuran Sampel			H	P
n ₁	n ₂	n ₃			n ₁	n ₂	n ₃		
2	1	1	2.70000	.500	4	3	2	6.4444	.008
								6.6000	.011
2	2	1	3.6000	.200				5.4444	.046
								5.4000	.051
2	2	2	4.5714					4.5111	.098
			3.7143					4.4444	.102
3	1	1	3.2000						
					4	3	3	6.7455	.010
3	2	1	4.2857					6.7091	.013
			3.8571					5.7909	.046
								5.72723	.050
3	2	2	5.3572					4.7091	.092
			4.7143					4.7000	.101
			4.5000						
			4.4643		4	4	1	6.6667	.010
								6.1667	.022
3	3	1	5.1429					4.9667	.048
			4.5714					4.8667	.054
			4.0000					4.1667	.082
								4.0667	.102
3	3	2	6.2500						
			5.3611		4	4	2	7.0364	.006

			5.1389				6.8727	.011	
			4.5556				5.4545	.046	
			4.2500				5.2364	.052	
							4.5545	.098	
3	3	3	7.2000				4.4455	.103	
			6.4889						
			5.6889		4	4	3	7.1439	.010
			5.6000					7.1364	.011
			5.0667					5.5985	.049
			4.6222					5.5758	.051
								4.5455	.099
4	1	1	3.5714					4.4773	.102
4	2	1	4.8214		4	4	4	7.6538	.008
			4.5000					7.5385	.011
			4.0179					5.6923	.049
								5.6538	.054
4	1	1	6.0000					4.6539	.097
			5.3333					4.5001	.104
			5.1250						
			4.4583		5	1	1	3.8571	.143
			4.1667						
					5	2	1	5.2500	.036
4	3	1	5.833					5.0000	.048
			5.2083					4.4500	.071
			5.0000					4.2000	.095
			3.8889					4.0500	.119
Ukuran Sampel			H	P	Ukuran Sampel			H	P
n1	n2	n3			n1	n2	n3		
5	2	2	6.5333	.008				5.6308	.050
			6.1333	.013				4.5487	.099
			5.1600	.034				4.5231	.103
			5.0400	.056					
			4.3733	.090	5	4	4	7.7604	.009

			4.2933	.122				7.7604	.011
								5.6571	.049
5	3	1	6.4000	.012				5.6176	.050
			4.9600	.048				4.6187	.100
			4.8711	.052				4.5527	.102
			4.0178	.095					
			3.8400	.123	5	5	1	7.3091	.009
									.011
5	3	2	6.9091						.046
			6.8218						.053
			5.2509						.086
			5.1055						.105
			4.6509						
			4.4945						.010
									.010
3	3	2	6.2500						
			5.3611		4	4	2	7.0364	.006
			5.1389					6.8727	.011
			4.5556					5.4545	.046
			4.2500					5.2364	.052
								4.5545	.098
3	3	3	7.2000					4.4455	.103
			6.4889						
			5.6889		4	4	3	7.1439	.010
			5.6000					7.1364	.011
			5.0667					5.5985	.049
			4.6222					5.5758	.051
								4.5455	.099
4	1	1	3.5714					4.4773	.102
4	2	1	4.8214		4	4	4	7.6538	.008
			4.5000					7.5385	.011
			4.0179					5.6923	.049
								5.6538	.054

4	1	1	6.0000					4.6539	.097
			5.3333					4.5001	.104
			5.1250						
			4.4583		5	1	1	3.8571	.143
			4.1667						
					5	2	1	5.2500	.036
4	3	1	5.833					5.0000	.048
			5.2083					4.4500	.071
			5.0000					4.2000	.095
			3.8889					4.0500	.119

BIODATA PENULIS

Dr. Nalom siagian, MM. Dilahirkan di Janjimaria 18 April 1966. Menempuh pendidikan tingkat tinggi sebagai berikut: Tahun 1990 lulus S-1 Program studi Administrasi Bisnis Univ. HKBP Nommensen Medan. Pada tahun 2006 lulus S-2 Program Magister Manajemen dari Univ.HKBP Nommensen Medan dan gelar doktor di raih dari Program Pasca Sarja Universitas Sumatera Utara pada tahun 2020. Selanjutnya, riwayat pekerjaan penulis adalah pada tahun 1990 menjadi dosen tidak tetap di Univ. HKBP Nommensen Medan, tahun tahun 1992 diangkat sebagai dosen tetap Yayasan Univ.HKBP Nommensen Medan sampai sekarang.

Untuk jabatan struktural, penulis pada tahun 1993 - 1996 menjabat sebagai Sekretaris Program Studi Administrasi Bisnis, Tahun 1996 - 1999 menjabat sebagai Deputy Manager Pusat Jasa Ketenagakerjaan (PJK) Universitas HKBP Nommensen Medan, Tahun 1999 - 2001 menjabat Kepala Perpustakaan Universitas HKBP Nommensen, Tahun 2001 - 2004 Wakil Dekan III FIIPOL Universitas HKBP Nommensen, Tahun 2004 - 2006 mendapat Tugas Belajar menempuh Program S-2, Tahun 2008 - 2011 dipercaya sebagai Sekretaris Pusat Kewirausahaan Mahasiswa Universitas HKBP Nommensen, Tahun 2011 - 2015 Menjabat Wakil Dekan II FISIPOL Universitas HKBP Nommensen, Tahun 2015 – 2020 mendapat tugas belajar menempuh Program Doktor, dan Tahun 2021 Wakil Dekan I FISIPOL UHN MEDAN.

Penulis mengampu mata kuliah antara lain: Statistik Sosial, Metode Penelitian Bisnis Akuntansi Biaya, Kewirausahaan, Riset Operasi Bisnis dan Analisis dan Evaluasi Bisnis. Penulis tinggal di Jl. Tangguk Damai Raya No. 215 Blok 2 Grya Martubung, Medan Labuhan.
Email : nalomsiagian278@gmail.com