PENGANTAR STATISTIKA UNTUK ILMU SOSIAL

Penulis:

Zul Fadli, S.E., M.A.P.
Yanti Murni, S.E., M.M.
Marsuddin Musa, S.Si., M.Stat.
Putri Anugrah Cahya Dewi, M.Pd.
Ketut Queena Fredlina, S.Si., M.Si.
Dr. Darham, S.Pd., M.P.d
Dr. H. Nanang, M.Pd.
Romi Mesra, S.Pd., M.Pd.
Hartina Husain, S.Si., M.Stat.

GET PRESS INDONESIA

PENGANTAR STATISTIKA UNTUK ILMU SOSIAL

Penulis :

Zul Fadli, S.E., M.A.P.
Yanti Murni, S.E., M.M.
Marsuddin Musa, S.Si., M.Stat.
Putri Anugrah Cahya Dewi, M.Pd.
Ketut Queena Fredlina, S.Si., M.Si.
Dr. Darham, S.Pd., M.P.d
Dr. H. Nanang, M.Pd.
Romi Mesra, S.Pd., M.Pd.
Hartina Husain, S.Si., M.Stat.

ISBN :
Editor : Nanny Mayasari, S.Pd., M.Pd., CQMS.
Penyunting: Yuliatri M.Hum.
Desain Sampul dan Tata Letak : Atyka, S.Pd.

Penerbit : Get Press Indonesia
Anggota IKAPI No. 033/SBA/2022

Redaksi :
Jl. Palarik Air Pacah RT 001 RW 006
Kelurahan Air Pacah Kecamatan Koto Tangah
Padang Sumatera Barat
Website : www.getpress.co.id
Email : globaleksekutifteknologi@gmail.com

Cetakan pertama, Oktober 2023

KATA PENGANTAR

Puji Syukur kami panjatkan kepada Allah SWT, berkat Rahmat dan hidayah-Nya, kami dapat menyelesaikan buku ini yang berjudul "**PENGANTAR STATISTIKA UNTUK ILMU SOSIAL".** Buku ini merupakan karya kolaboratif sebagai panduan komprehensif untuk memahami dan menguasai konsep-konsep statistika merupakan alat yang kuat dalam menjelaskan fenomena sosial dan mengaplikasikannya.

Buku ini membahas tentang pengantar dan Sejarah perkembangan statistika sebagai ilmu pengetahuan, yang telah memberikan landasan kuat bagi ststistikaa dalam berbagai disiplin ilmu. Skala pengukuran memberikan pemahaman jenis-jenis skala pengukuran, termasuk skala nominal, ordinal, interval, dan rasio. Bagaimana memilih skala yang sesuai untuk data penelitian ilmu sosial. Selain itu, konsep distribusi data dan pentingnya bilanganbaku dalam ststistika. Bagaimana menghitungan baku untuk mengevaluasi sebaran data dan membahas konsep peluang, distribusi probalitas yang dapat digunakan dalam pengambilan sampel dari populasi dan inferensi tentang populasi.

Statistika inferensial merupakan alat penting untuk membuat kesimpulan berdasarkan data sampel. Bab ini, mempelajari pengujian hipotesis, pengambilan sampel, dan teknik-teknik yang digunakan untuk membuat kesimpulan yang lebih luas tentang populasi berdasarkan sampel yang ada. Uji korelasi dan asosiasi membantu mengidentifikasi dan mengukur hubungan antar variabel kategorik. Bab terakhir, berbagai jenis korelasi, penggunaan analisis data, dan bagaimana hasilnya dapat diinterprestasikan dalam ilmu sosial.

Buku ini cocok untuk mahasiswa S1 dan S2, akademisi, dan praktisi dalam memahami dan mengembangkan penelitian ilmu sosial. Semoga buku ini memberikan banyak manfaat dan pahala amal jariah untuk para penulisnya.

Padang, Oktober 2023
Editor
Nanny Mayasari, S.Pd., M.Pd., CQMS.

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB 1

PENGANTAR DAN SEJARAH STATISTIKA

Oleh Zul Fadli, S.E., M.A.P.

1.1. Pendahuluan

Statistika merupakan ilmu yang kuat dan sangat diperlukan dalam mengumpulkan, menganalisis, dan menginterpretasikan data sosial, sehingga memungkinkan kita untuk memahami berbagai fenomena sosial dengan lebih akurat dan obyektif (Widodo and Andawaningtyas, 2017). Dalam ilmu sosial, pengumpulan data dan analisis statistik menjadi hal yang sangat vital. Data sosial mencakup beragam aspek kehidupan manusia, seperti tingkah laku, sikap, pendapat, ekonomi, politik, pendidikan, dan banyak lagi. Melalui statistika, kita dapat menemukan wawasan penting tentang dinamika masyarakat, interaksi manusia, dan faktor-faktor yang mempengaruhi perilaku individu maupun kelompok. Pentingnya statistika dalam ilmu sosial semakin meningkat seiring dengan berkembangnya era digital dan teknologi informasi. Data sosial yang melimpah dan kompleks memerlukan pendekatan analitis yang tepat untuk menggali informasi yang berarti dan relevan bagi pengambilan keputusan.

Statistika dalam ilmu sosial sangatlah penting karena dapat memberikan kerangka kerja dan alat analisis yang krusial dalam menghadapi data sosial yang kompleks dan beragam. Beberapa aspek yang melatarbelakangi pentingnya statistika dalam ilmu sosial adalah sebagai berikut:

1. Ilmu sosial mencakup beragam fenomena sosial yang melibatkan manusia sebagai individu atau dalam kelompok.

Data-data sosial seringkali lebih kompleks daripada data dalam ilmu alam atau eksakta. Statistika memberikan metode dan teknik untuk mengumpulkan, menyusun, dan merapikan data yang beragam, sehingga memungkinkan peneliti untuk mendapatkan pemahaman yang lebih komprehensif tentang fenomena sosial.

2. Statistika membantu ilmu sosial dalam menganalisis data dan mengambil kesimpulan berdasarkan bukti empiris. Dengan menggunakan teknik statistik, peneliti dapat mengidentifikasi hubungan antara variabel, menguji hipotesis, dan menggeneralisasi hasil penelitian kepada populasi yang lebih luas. Tanpa statistika, analisis data sosial akan kurang obyektif dan kurang mendalam.
3. Statistika berperan penting dalam validasi temuan penelitian ilmu sosial. Melalui analisis statistik yang tepat, peneliti dapat mengukur tingkat kepercayaan terhadap hasil penelitian, sehingga temuan dapat diandalkan dan diyakini. Validitas statistik juga membantu menghindari bias dan kesalahan dalam interpretasi data.
4. Statistika memberikan landasan yang kuat dalam perencanaan dan pengambilan kebijakan di berbagai bidang ilmu sosial. Melalui analisis data statistik, pemangku kepentingan dapat mendapatkan gambaran yang lebih jelas tentang situasi dan kondisi masyarakat, sehingga kebijakan yang diambil dapat lebih tepat dan relevan.
5. Dalam era digital dan big data saat ini, data sosial semakin melimpah dan kompleks. Statistika menjadi semakin penting dalam menghadapi tantangan ini, untuk mengolah data yang besar dan beragam, serta menggali informasi berharga dari data tersebut (Widodo and Andawaningtyas, 2017).

Secara keseluruhan, latar belakang statistika dalam ilmu sosial mencerminkan kebutuhan dan relevansinya dalam menghadapi fenomena sosial yang kompleks dan beragam. Dengan statistika, ilmu sosial menjadi lebih kuat dalam menghasilkan temuan yang obyektif, mendalam, dan berbasis

bukti-bukti empiris, sehingga kontribusinya dalam pemahaman dan pemecahan masalah sosial semakin meningkat.

1.2. Definisi Statistika

Statistika merupakan cabang ilmu pengetahuan yang berkaitan dengan pengumpulan, pengolahan, interpretasi, dan analisis data (Wulandari, 2022). Tujuan utama statistika adalah untuk menyajikan data secara terstruktur, memahami pola dan hubungan dalam data, dan mengambil kesimpulan atau membuat prediksi berdasarkan data yang dianalisis. Berikut beberapa definisi statistika menurut para ahli:

1. Menurut David R. Anderson, Dennis J. Sweeney, Thomas A. Williams dalam buku yang berjudul "Statistics for Business and Economics" mengemukakan bahwa statistika adalah ilmu tentang pengumpulan, analisis, interpretasi, presentasi, dan pengambilan keputusan berdasarkan data.
2. Menurut Gerald Keller dalam buku yang berjudul "Statistics for Management and Economics" mengatakan bahwa statistika adalah ilmu yang berkaitan dengan pengumpulan, pengolahan, interpretasi, analisis, dan presentasi data.
3. Menurut Richard A. Johnson, Gouri K. Bhattacharyya dalam buku yang berjudul "Statistics: Principles and Methods" mengatakan bahwa statistika adalah ilmu yang mempelajari cara mengumpulkan, mengorganisir, menganalisis, dan menginterpretasikan data untuk membuat kesimpulan yang berdasarkan bukti.
4. Menurut Paul Newbold, William L. Carlson, Betty Thorne dalam buku yang berjudul "Statistics for Business and Economics" mengemukakan bahwa statistika merupakan ilmu yang berhubungan dengan pengumpulan, analisis, interpretasi, presentasi, dan penggunaan data untuk membuat keputusan.
5. Menurut Richard D. De Veaux, Paul F. Velleman, David E. Bock dalam buku yang berjudul "Intro Stats" mengatakan bahwa statistika adalah ilmu yang mempelajari cara pengumpulan, analisis, interpretasi, dan presentasi data

untuk membantu dalam pengambilan keputusan (Wulandari, 2022).

Secara umum, para ahli sepakat bahwa statistika melibatkan pengumpulan dan analisis data serta penggunaannya dalam pengambilan keputusan. Statistika memainkan peran penting dalam berbagai bidang, termasuk bisnis, ekonomi, sains, kedokteran, sosial, dan banyak lagi.

1.3. Sejarah Statistika

Statistika telah ada sejak zaman kuno, meskipun mungkin tidak disadari sebagai ilmu yang terstruktur pada saat itu. Pada masa lalu, orang-orang menggunakan metode pengumpulan data dan analisis sederhana untuk membantu pengambilan keputusan dalam berbagai aspek kehidupan, seperti pertanian, perdagangan, dan perencanaan sosial. Contoh awal penggunaan statistika dapat ditemukan dalam perhitungan penduduk, produksi pertanian, dan pengukuran astronomi di Mesir Kuno, Tiongkok Kuno, dan daerah-daerah lain di dunia (Nisfiannoor, 2009). Namun, pengembangan formal statistika sebagai ilmu dimulai pada abad ke-17 dan ke-18. Salah satu tonggak penting dalam sejarah statistika adalah penemuan hukum probabilitas oleh matematikawan Prancis, Blaise Pascal dan Pierre de Fermat. Hukum probabilitas memberikan landasan teoretis bagi konsep peluang, yang menjadi dasar penting dalam analisis statistika.

1.3.1. Sejarah Statistika di Dunia

Sejarah statistika di dunia memiliki akar yang kuat dan meluas sepanjang waktu. Berikut adalah rangkuman yang detail mengenai sejarah statistika di dunia:

1. Awal Mula dan Zaman Kuno
   Praktik pengumpulan dan analisis data telah ada sejak zaman kuno di berbagai peradaban seperti Mesir Kuno, Tiongkok Kuno, dan Babilonia. Contoh awal penggunaan statistika dapat ditemukan dalam perhitungan penduduk, produksi pertanian, dan pengukuran astronomi. Pada periode ini,

praktik statistika umumnya didasarkan pada pengamatan langsung, hitungan sederhana, dan perbandingan jumlah.

2. Perkembangan Awal
   Pada abad ke-17, konsep peluang dan hukum probabilitas muncul sebagai dasar statistika modern. Matematikawan Prancis Blaise Pascal dan Pierre de Fermat membuat kontribusi signifikan dalam pengembangan hukum probabilitas. Pada waktu itu, statistika masih terbatas pada penggunaan peluang dalam permainan judi dan perhitungan keuangan.
3. Zaman Pencerahan
   Pada abad ke-18, statistika berkembang pesat dengan kontribusi dari berbagai ilmuwan dan matematikawan. Penyair dan matematikawan Inggris, Thomas Bayes, memberikan kontribusi penting dalam statistika dengan teorema Bayes yang memungkinkan perhitungan probabilitas posterior berdasarkan informasi sebelumnya. Statistika mulai diterapkan dalam bidang sosial dan ekonomi, dan metode analisis data yang lebih sistematis dikembangkan.
4. Perkembangan Metode Statistik
   Pada abad ke-19, statistika menjadi lebih terstruktur dan metode statistik yang lebih maju dikembangkan. Carl Friedrich Gauss, matematikawan Jerman, memainkan peran penting dalam pengembangan metode statistik dan teori distribusi normal. Statistik mulai diterapkan dalam bidang ilmu alam, seperti astronomi, fisika, dan biologi. Statistik menjadi alat yang penting dalam analisis populasi, penelitian sosial, dan analisis ekonomi.
5. Era Modern
   Pada awal abad ke-20, statistika mengalami perkembangan pesat dengan kemajuan teknologi dan komputerisasi. Statistik terus diterapkan dalam berbagai disiplin ilmu, termasuk kedokteran, sains lingkungan, teknik, dan ilmu sosial. Konsep inferensi statistik dan uji hipotesis menjadi lebih maju, memungkinkan penarikan kesimpulan dari sampel ke populasi secara lebih akurat. Perkembangan komputer memungkinkan analisis data yang lebih kompleks

dan memperluas peran statistika dalam pengambilan keputusan.

6. Statistika pada Abad ke-21
   Dalam era digital dan big data, statistika menjadi semakin penting. Analisis data besar (big data) dan pembelajaran mesin (machine learning) menjadi fokus utama dalam statistika saat ini. Statistik juga digunakan dalam pengembangan model prediktif, pengoptimalan, dan pengambilan keputusan berbasis data. Konsep statistika seperti regresi, analisis varians, dan pengujian hipotesis tetap menjadi bagian penting dalam metodologi statistik (Nisfiannoor, 2009).

Seiring dengan perkembangan zaman, statistika terus beradaptasi dan berkembang. Penerapan statistika semakin meluas dalam berbagai bidang, dan metodologi statistik terus berkembang untuk mengatasi tantangan analisis data yang semakin kompleks. Sejarah statistika memberikan dasar yang kuat bagi pemahaman dan penerapan statistika modern di dunia saat ini.

1.3.2. Sejarah Statistik di Indonesia

Sejarah statistika di Indonesia mencerminkan perkembangan dan penggunaan data dalam menghadapi berbagai tantangan dan kebutuhan pembangunan negara. Berikut adalah penjelasan rinci tentang sejarah statistika di Indonesia:

1. Era Kolonial
   Pada masa penjajahan, pengumpulan data statistik pertama di Indonesia dilakukan oleh pemerintah kolonial Belanda. Tujuan utama pengumpulan data tersebut adalah untuk mengelola sumber daya alam dan melacak perkembangan ekonomi koloni. Data yang dikumpulkan meliputi informasi tentang penduduk, hasil pertanian, produksi industri, dan perdagangan.
2. Masa Kemerdekaan

Setelah kemerdekaan Indonesia, statistika menjadi lebih penting dalam perencanaan pembangunan nasional. Pada tahun 1950, didirikan Badan Pusat Statistik (BPS) sebagai lembaga pemerintah yang bertanggung jawab untuk mengumpulkan, menganalisis, dan menyajikan data statistik di Indonesia. BPS menjadi lembaga sentral dalam menghasilkan statistik resmi yang digunakan dalam perencanaan dan pengambilan keputusan pemerintah.

3. Pengembangan Infrastruktur Statistik
   Sejak tahun 1950-an, Indonesia mengembangkan infrastruktur statistik yang lebih baik untuk mengumpulkan data yang lebih luas dan akurat. BPS memperluas jaringan survei dan pemetaan statistik di seluruh Indonesia. Penyelenggaraan sensus penduduk, sensus pertanian, dan survei sosial-ekonomi menjadi rutin dilakukan untuk mendapatkan informasi yang diperlukan dalam perencanaan pembangunan.
4. Perkembangan Metode dan Teknik Statistik
   Pada tahun 1960-an, terjadi peningkatan dalam penggunaan metode dan teknik statistik di Indonesia. Pelatihan statistik diberikan kepada personel BPS dan perencana pembangunan, serta statistik mulai diterapkan dalam penelitian sosial, ekonomi, dan ilmiah. Pada era ini, metode statistik seperti regresi, analisis multivariat, dan analisis deret waktu mulai diperkenalkan dan diterapkan dalam konteks Indonesia.
5. Era Digital dan Transformasi Statistik
   Dalam beberapa dekade terakhir, perkembangan teknologi informasi dan komunikasi telah mengubah cara statistik dilakukan di Indonesia. Penggunaan komputer dan teknologi digital memungkinkan pengolahan data yang lebih cepat dan akurat. BPS mengembangkan sistem informasi statistik online yang memudahkan akses publik terhadap data dan statistik yang diperlukan. Big data dan analisis data juga menjadi fokus dalam menghadapi tantangan pengolahan data yang semakin besar dan kompleks (Irianto, 2021).

Pada saat ini, statistika terus berkembang dan menjadi bagian penting dari perencanaan pembangunan, pengambilan keputusan, dan penelitian di Indonesia. BPS terus berperan sebagai lembaga utama dalam menghasilkan dan menyebarkan data statistik yang akurat dan dapat diandalkan bagi masyarakat, sektor publik, dan sektor swasta.

1.4. Hubungan Statistika dan Ilmu Sosial

Statistika memiliki hubungan yang erat dengan ilmu sosial dan berperan sebagai alat yang penting dalam analisis data dan pengambilan keputusan dalam konteks ilmu sosial. Statistika dapat membantu dalam pengumpulan data yang sistematis dan terstruktur dalam ilmu sosial. Melalui teknik sampling dan survei, statistika membantu ilmu sosial dalam mendapatkan data yang representatif tentang populasi atau kelompok tertentu yang sedang diteliti (Ismail, 2018). Statistika juga membantu dalam menganalisis, menggambarkan, dan menyajikan data secara terperinci. Dengan menggunakan tabel, grafik, dan ukuran statistik seperti mean, median, dan modus, statistika membantu menggambarkan distribusi data dan memvisualisasikan pola-pola yang ada. Selanjutnya statistika memberikan metode dan teknik analisis yang diperlukan untuk menguji hipotesis dan menjawab pertanyaan-pertanyaan penelitian dalam ilmu sosial. Contohnya, analisis regresi digunakan untuk mengevaluasi hubungan antara variabel-variabel dalam ilmu politik, sosiologi, dan ekonomi.

Statistika dapat digunakan untuk membangun model dan prediksi dalam ilmu sosial. Dengan menggunakan teknik seperti analisis regresi, analisis deret waktu, dan pembelajaran mesin, statistika dapat membantu dalam membangun model yang dapat memprediksi perilaku atau peristiwa masa depan berdasarkan pola yang terlihat dalam data historis (Ismail, 2018). Sebagai alat evaluasi yang kuat, statistika dapat mengukur dampak kebijakan dan program sosial. Dengan menggunakan metode eksperimen atau quasi-eksperimen, statistika membantu dalam menentukan efektivitas suatu kebijakan atau program dalam mencapai tujuan sosial yang diinginkan.

Kemudian dalam kategori inferensi statistik memungkinkan ilmu sosial untuk membuat generalisasi atau kesimpulan lebih luas berdasarkan data yang terbatas. Melalui metode seperti interval kepercayaan dan uji hipotesis, statistika membantu mengukur tingkat ketidakpastian dan mengambil kesimpulan yang berdasarkan bukti data (Sutopo and Slamet, 2017). Secara keseluruhan, statistika menjadi penting dalam ilmu sosial karena membantu dalam mengumpulkan, menganalisis, dan menginterpretasikan data, serta memberikan landasan untuk pengambilan keputusan berdasarkan bukti-bukti empiris. Dalam era digital dan big data saat ini, statistika semakin relevan dalam menganalisis dan memahami fenomena sosial yang kompleks.

1.5. Fungsi, Tujuan dan Manfaat Statistika Ilmu Sosial

1.5.1. Fungsi Statistika dalam Ilmu Sosial

Berikut fungsi statistika dalam ilmu sosial:

1. Statistika membantu dalam pengumpulan data yang sistematis dan representatif tentang fenomena sosial yang sedang diteliti. Melalui teknik sampling dan survei, statistika memastikan bahwa data yang dikumpulkan mencerminkan populasi atau kelompok tertentu dengan akurat.
2. Statistika memberikan metode dan teknik analisis yang diperlukan untuk mengolah data sosial yang kompleks. Statistika memungkinkan peneliti untuk mengidentifikasi pola, hubungan, perbedaan, dan variabilitas dalam data sosial. Ini membantu dalam menjawab pertanyaan penelitian dan mengungkapkan wawasan tentang fenomena sosial yang sedang diteliti.
3. Statistika membantu dalam memahami dan menginterpretasikan data sosial. Dengan menggunakan ukuran statistik seperti mean, median, dan modus, statistika memungkinkan peneliti untuk memberikan deskripsi numerik tentang data. Statistika juga membantu dalam memvisualisasikan data melalui grafik dan diagram, yang mempermudah pemahaman dan komunikasi temuan

penelitian kepada pembaca atau pemangku kepentingan (Yusi and Idris, 2020).

1.5.2. Tujuan Statistika dalam Ilmu Sosial

Berikut tujuan statistika dalam ilmu sosial:

1. Memberikan deskripsi yang sistematis tentang fenomena sosial yang sedang diteliti. Statistika membantu dalam menggambarkan karakteristik dan distribusi data sosial, sehingga memungkinkan peneliti untuk memahami dan menjelaskan fenomena tersebut secara kuantitatif.
2. Menjelaskan hubungan dan pola yang ada dalam data sosial. Melalui analisis statistik, peneliti dapat mengidentifikasi korelasi, asosiasi, atau hubungan sebab-akibat antara variabel sosial. Ini membantu dalam memahami faktor-faktor yang mempengaruhi fenomena sosial dan memperoleh wawasan yang lebih dalam tentang kompleksitas sosial.
3. Menguji hipotesis atau pernyataan yang diajukan dalam penelitian. Statistika menyediakan alat dan teknik untuk menguji keberartian statistik dan mengambil kesimpulan berdasarkan bukti-bukti empiris yang ditemukan dalam data sosial (Yusi and Idris, 2020).

1.5.3. Manfaat Statistika dalam Ilmu Sosial

Berikut manfaat statistika dalam ilmu sosial:

1. Membantu dalam menjaga objektivitas dalam analisis data sosial. Dengan menggunakan metode statistik yang baku, peneliti dapat menghindari bias dan asumsi pribadi yang dapat memengaruhi interpretasi dan kesimpulan.
2. Memberikan metode yang kuat untuk menguji keandalan data dan hasil penelitian sosial. Melalui teknik statistik seperti uji signifikansi dan interval kepercayaan, peneliti dapat mengukur tingkat ketidakpastian dan memastikan bahwa hasil penelitian memiliki dasar empiris yang kuat.
3. Memungkinkan peneliti untuk membuat generalisasi tentang fenomena sosial yang sedang diteliti. Dengan

menggunakan teknik sampling yang tepat dan analisis statistik yang benar, peneliti dapat membuat inferensi yang valid tentang populasi yang lebih luas berdasarkan data yang terbatas.

4. Membantu dalam pengambilan keputusan yang lebih informan dan rasional dalam konteks ilmu sosial. Dengan menggunakan bukti-bukti data yang dikumpulkan dan dianalisis secara statistik, pengambilan keputusan dapat didukung oleh informasi yang lebih akurat dan obyektif (Yusi and Idris, 2020).

Dengan demikian, statistika memiliki peran yang signifikan dalam ilmu sosial. Melalui pengumpulan data yang akurat, analisis yang tepat, dan interpretasi yang obyektif, statistika memberikan landasan yang kuat untuk memahami dan menjelaskan fenomena sosial dalam konteks ilmu sosial.

1.6. Studi Kasus Ilmu Sosial dengan Statistika

Salah satu contoh studi kasus ilmu sosial yang menggunakan statistika adalah penelitian mengenai faktor-faktor yang mempengaruhi kepuasan kerja karyawan di suatu perusahaan. Dalam penelitian ini, statistika digunakan untuk menganalisis data yang dikumpulkan melalui survei untuk mengidentifikasi hubungan antara variabel-variabel yang relevan. Misalnya, peneliti dapat mengumpulkan data mengenai faktor-faktor seperti tingkat gaji, lingkungan kerja, peluang pengembangan karir, dan dukungan dari atasan. Setelah data terkumpul, statistika dapat digunakan untuk menganalisis data tersebut dan menjawab pertanyaan penelitian, seperti:

1. Apakah ada hubungan antara tingkat gaji dan kepuasan kerja?
2. Apakah ada perbedaan kepuasan kerja antara karyawan yang mendapatkan dukungan dari atasan dengan yang tidak?
3. Apakah ada hubungan antara peluang pengembangan karir dan kepuasan kerja?

Statistika dapat memberikan alat dan teknik untuk menjawab pertanyaan-pertanyaan tersebut. Misalnya, analisis regresi dapat digunakan untuk mengidentifikasi hubungan antara variabel-variabel tersebut. Selain itu, uji hipotesis statistik juga dapat digunakan untuk menguji signifikansi hubungan antara variabel-variabel tersebut. Dengan menggunakan metode statistik, peneliti dapat menganalisis data yang terkumpul dengan lebih obyektif dan mendapatkan wawasan yang lebih mendalam tentang faktor-faktor yang berkontribusi terhadap kepuasan kerja karyawan. Hasil dari analisis statistik ini dapat memberikan dasar yang kuat untuk pengambilan keputusan yang lebih baik dalam upaya meningkatkan kepuasan kerja dan kesejahteraan karyawan di perusahaan tersebut (Lubis, 2021).

Selain penelitian mengenai kepuasan kerja, statistika juga digunakan dalam berbagai studi kasus ilmu sosial lainnya. Contoh lainnya adalah penelitian mengenai faktor-faktor yang mempengaruhi tingkat partisipasi masyarakat dalam kegiatan sosial atau politik. Misalnya, seorang peneliti ingin mengetahui apakah tingkat pendidikan, usia, dan tingkat penghasilan mempengaruhi partisipasi masyarakat dalam kegiatan sosial seperti kegiatan sukarela atau partisipasi politik seperti pemilihan umum. Peneliti dapat mengumpulkan data melalui survei atau wawancara terstruktur dan menggunakan statistika untuk menganalisis data tersebut. Dalam hal ini, statistika dapat digunakan untuk menjawab pertanyaan-pertanyaan seperti:

1. Apakah ada hubungan antara tingkat pendidikan dan partisipasi masyarakat dalam kegiatan sosial atau politik?
2. Apakah ada perbedaan tingkat partisipasi masyarakat antara kelompok usia yang berbeda?
3. Apakah ada hubungan antara tingkat penghasilan dan partisipasi masyarakat dalam kegiatan sosial atau politik?

Dengan menggunakan metode statistik, peneliti dapat menganalisis data dan mengidentifikasi hubungan antara variabel-variabel tersebut. Misalnya, uji korelasi dapat digunakan untuk melihat hubungan antara tingkat pendidikan dan partisipasi masyarakat. Selain itu, analisis statistik seperti uji

beda atau uji regresi juga dapat digunakan untuk menguji hubungan antara variabel-variabel tersebut (Vivi Silvia, 2020). Hasil analisis statistik tersebut dapat memberikan pemahaman yang lebih mendalam tentang faktor-faktor yang mempengaruhi tingkat partisipasi masyarakat dalam kegiatan sosial atau politik. Dalam konteks ini, statistika berfungsi untuk memvalidasi hipotesis, mengidentifikasi variabel yang paling signifikan, dan memberikan dasar empiris untuk pengambilan keputusan dan pengembangan kebijakan yang lebih baik dalam rangka meningkatkan partisipasi masyarakat.

Studi kasus ini mencerminkan bagaimana statistika digunakan sebagai alat yang penting dalam penelitian ilmu sosial untuk menggali hubungan antara variabel-variabel yang kompleks dan memahami fenomena sosial yang melibatkan banyak faktor. Secara keseluruhan, statistika memainkan peran yang krusial dalam studi kasus ilmu sosial. Melalui pengumpulan dan analisis data dengan menggunakan metode statistik, peneliti dapat mengungkap hubungan dan pola yang ada, memberikan dasar empiris untuk pengambilan keputusan, dan meningkatkan pemahaman tentang fenomena sosial yang kompleks.

1.7. Penutup

Statistika memainkan peran yang sangat penting dalam ilmu sosial. Fungsi utamanya adalah untuk mengumpulkan, menganalisis, dan menginterpretasikan data sosial secara sistematis. Tujuan statistika dalam ilmu sosial adalah untuk mendeskripsikan fenomena sosial, menjelaskan hubungan dan pola yang ada, serta menguji hipotesis yang diajukan dalam penelitian. Manfaat statistika dalam ilmu sosial meliputi objektivitas analisis, keandalan hasil penelitian, kemampuan untuk melakukan generalisasi, dan dukungan dalam pengambilan keputusan yang informan. Dengan menggunakan metode statistik, ilmu sosial dapat menghasilkan pengetahuan yang lebih teruji secara empiris, mengungkapkan pola-pola yang ada dalam fenomena sosial, dan mengambil keputusan yang lebih baik dalam berbagai konteks. Statistika membantu mengatasi bias dan

asumsi pribadi, serta memberikan landasan yang kuat untuk mengambil kesimpulan yang didukung oleh data yang obyektif.

Dalam era informasi dan pengumpulan data yang melimpah, kemampuan untuk memahami, menganalisis, dan menginterpretasikan data dengan benar menjadi semakin penting dalam ilmu sosial. Dengan bantuan statistika, peneliti dapat menghadapi tantangan ini dan memperoleh pemahaman yang lebih mendalam tentang fenomena sosial yang kompleks.

DAFTAR PUSTAKA

Irianto, H.A. (2021), *Statistik Untuk Ilmu Sosial: Aplikatif Untuk Ilmu-Ilmu Sosial*, Prenada Media.

Ismail, H.F. (2018), *Statistika Untuk Penelitian Pendidikan Dan Ilmu-Ilmu Sosial*, Kencana.

Lubis, Z. (2021), *Statistika Terapan Untuk Ilmu-Ilmu Sosial Dan Ekonomi*, Penerbit Andi.

Nisfiannoor, M. (2009), *Pendekatan Statististika Modern Untuk Ilmu Sosial*, Penerbit Salemba.

Sutopo, E.Y. and Slamet, A. (2017), *Statistik Inferensial*, Penerbit Andi.

Vivi Silvia, S.E. (2020), *Statistika Deskriptif*, Penerbit Andi.

Widodo, A. and Andawaningtyas, K. (2017), *Pengantar Statistika*, Universitas Brawijaya Press.

Wulandari, O.A.D. (2022), *Statistika Untuk Ilmu Sosial: Teori Dan Aplikatif Untuk Ilmu–Ilmu Sosial*, Vol. 1, Zahira Media Publisher.

Yusi, M.S. and Idris, U. (2020), *Statistika Untuk Ekonomi, Bisnis, & Sosial*, Penerbit Andi.

BAB 2

SKALA PENGUKURAN

Oleh Yanti Murni, S.E., M.M.

2.1. Pendahuluan

Skala pengukuran adalah sistem yang digunakan untuk menggambarkan sifat dan tingkat data yang diukur dalam penelitian atau pengumpulan data. Skala pengukuran menentukan jenis informasi yang dapat diperoleh, metode analisis yang sesuai yang dapat digunakan untuk memahami dan menggambarkan data tersebut (Junaidi, 2015).

Ada empat jenis skala pengukuran yang umum digunakan dalam statistika, yaitu:

1. **Skala Nominal**. Merupakan skala yang digunakan untuk mengkategorikan data ke dalam kelompok atau kategori yang berbeda, tanpa memiliki urutan atau tingkatan.
2. **Skala Ordinal**. Skala yang memungkinkan data diurutkan dalam urutan peringkat, tetapi tidak menunjukkan jarak atau selisih antara nilai-nilai data.
3. **Skala Interval**. Skala yang memiliki tingkatan peringkat dan jarak antara nilai-nilai data, tetapi tidak memiliki titik nol mutlak.
4. **Skala Rasio**. Skala yang memiliki peringkat, jarak dan titik nol mutlak yang mempunyai makna.

Sebab pentingnya skala pengukuran dalam penelitian ilmu sosial antara lain:

1. **Memberikan Struktur Data.** Skala pengukuran membantu memberikan struktur data yang jelas dan terorganisir,

sehingga memudahkan peneliti dalam mengumpulkan dan menganalisis data dengan tepat.

2. **Menentukan Metode Analisis.** Jenis skala pengukuran yang digunakan akan mempengaruhi metode analisis statistik yang dapat digunakan. Misalnya analisis regresi hanya dapat dilakukan pada data yang diukur dalam skala interval atau rasio.
3. **Menentukan Interpretasi Hasil.** Skala pengukuran menentukan interpretasi hasil penelitian. Misalnya, data yang diukur dalam skala nominal hanya dapat digunakan untuk mengidentifikasi perbedaan kategori, sementara data dalam skala rasio memungkinkan perbandingan proporsional.
4. **Validitas dan Reliabilitas.** Dalam penelitian ilmu sosial, penting untuk memastikan validitas dan reliabilitas data. Pemilihan skala pengukuran yang sesuai dapat membantu memastikan kualitas data dan validitas kesimpulan penelitian.
5. **Generalisasi.** Skala pengukuran yang benar membantu memungkinkan generalisasi hasil penelitian ke populasi yang lebih luas.

2.2. Skala Nominal

Skala nominal adalah salah satu jenis skala pengukuran dalam statistika yang digunakan untuk mengkategorikan atau mengklasifikasikan data ke dalam kelompok-kelompok atau kategori-kategori yang berbeda. Skala nominal hanya menyediakan informasi tentang perbedaan kualitatif antara kategori, tetapi tidak menyediakan informasi tentang urutan atau tingkatan antar kategori tersebut (Ilmiyah, 2013). Ciri utama skala nominal adalah bahwa tidak ada peringkat yang terukur antara nilai-nilai kategorikalnya. Pengukuran pada skala nominal hanya dapat mengidentifikasi perbedaan dalam atribut atau karakteristik, tetapi tidak ada perbandingan yang berarti di antara nilai-nilai data. Kategori dalam skala nominal bersifat saling eksklusif dan bersifat setara.

2.2.1. Contoh Skala Nominal

Beberapa contoh variabel yang termasuk skala nominal adalah sebagai berikut :

a. Jenis Kelamin. Misalnya kategori Laki-laki dan Perempuan.
b. Status Perkawinan. Misalnya kategori Belum Menikah, Menikah, Cerai dan Duda/Janda.
c. Agama. Misalnya kategori Islam, Kristen, Hindu, Buddha dan Konghucu.
d. Warna Rambut. Misalnya kategori Cokelat, Hitam, Pirang, Merah dan Abu-abu.

Dalam analisis data, untuk data yang diukur dalam skala nominal, hanya dapat dilakukan penghitungan frekuensi (jumlah kemunculan setiap kategori) dan persentase dari masing-masing kategori. Jenis analisis lain yang dapat digunakan untuk data nominal adalah uji chi-square untuk mengevaluasi hubungan antara variabel nominal yang berbeda (McHugh, 2012). Perlu diingat bahwa pada skala nominal, label atau nama kategori tidak memiliki nilai angka yang sebenarnya. Misal label "Laki-laki" atau "Perempuan" dalam variabel jenis kelamin tidak memiliki nilai numerik yang bermakna. Oleh karena itu, penggunaan operasi matematika seperti penjumlahan atau pengurangan tidak berlaku untuk data dalam skala nominal.

2.2.2. Karakteristik dan Sifat Data Skala Nominal

Karakteristik dan sifat data skala nominal adalah sebagai berikut:

1. Kategorikal.
   Data pada skala nominal terdiri dari kategori-kategori atau kelompok-kelompok diskrit yang saling eksklusif. Setiap unit data harus masuk ke dalam salah satu kategori dan tidak boleh termasuk dalam lebih dari satu kategori.
2. Tidak Berurutan.
   Data pada skala nominal tidak memiliki urutan atau tingkatan yang terukur di antara kategori-kategori. Artinya, kita tidak dapat menyusun data dalam urutan yang bermakna berdasarkan nilai nominalnya.

3. Identifikasi dan Penggolongan.
   Skala nominal digunakan untuk mengidentifikasi dan mengklasifikasikan unit data ke dalam kelompok-kelompok atau kategori-kategori berdasarkan atribut atau karakteristik tertentu. Kategori-kategori ini berfungsi sebagai label atau nama yang membantu mengidentifikasi kelompok tersebut.
4. Pengukuran Kualitatif.
   Variabel dalam skala nominal tidak dapat diukur dalam nilai numerik.
5. Operasi yang Terbatas.
   Penggunaan operasi matematika seperti penjumlahan, pengurangan dan perbandingan tidak berlaku untuk data dalam skala nominal. Kita tidak dapat melakukan operasi matematika pada label atau nama kategori.
6. Hanya Frekuensi dan Persentase.
   Analisis yang dapat dilakukan pada data dalam skala nominal adalah menghitung frekuensi (jumlah kemunculan setiap kategori) dan persentase dari masing-masing kategori. Misal kita dapat menghitung berapa banyak responden yang termasuk dalam kategori tertentu dan menyajikan hasilnya dalam bentuk persentase dari total responden.

2.2.3. Mengidentifikasi dan Kategorisasi Data Skala Nominal

Mengidentifikasi dan menggolongkan data ke dalam kategori skala nominal cukup sederhana karena data pada skala ini terdiri dari kategori-kategori diskrit yang bersifat saling eksklusif. Berikut adalah langkah - langkah untuk mengidentifikasi dan menggolongkan data ke dalam kategori pada skala nominal:

1. Identifikasi Kategori.
   Tentukan variabel apa yang akan diukur dalam penelitian atau pengumpulan data. Pastikan variabel tersebut sesuai dengan skala nominal, seperti jenis kelamin, agama, status perkawinan atau warna rambut.
2. Buat Daftar Kategori.

Buat daftar lengkap dari semua kategori yang mungkin muncul dalam data. Pastikan untuk mencakup semua kemungkinan kategori yang relevan dengan variabel yang diukur.

3. Kode Kategori.
   Setiap kategori dalam data harus diidentifikasi dengan kode unik yang membedakannya dari kategori lain. Kode ini dapat berupa angka atau huruf dan digunakan untuk mewakili setiap kategori dalam proses analisis.
4. Input Data.
   Inputkan data dari unit sampel (individu atau entitas) ke dalam database atau lembar kerja. Pastikan bahwa data sesuai dengan kategori yang tepat, sesuai dengan definisi kategori yang telah ditentukan sebelumnya.
5. Periksa Konsistensi.
   Pastikan data telah dimasukkan dengan benar dan konsisten dengan kategori yang ada. Periksa untuk memastikan tidak ada data yang masuk ke dalam lebih dari satu kategori dan setiap unit sampel dimasukkan ke dalam kategori yang paling sesuai.
6. Pelaporan Data.
   Saat melaporkan hasil penelitian atau analisis, tentukan kategori yang digunakan untuk setiap variabel nominal dan sertakan frekuensi atau persentase dari masing-masing kategori. Hal ini akan membantu pembaca memahami distribusi data dan karakteristik populasi yang diteliti.

Contoh:
Misalkan Kita melakukan survei tentang warna rambut di antara sekelompok orang. Variabel "Warna Rambut" merupakan variabel dalam skala nominal. Kita telah mendefinisikan kategori yang mungkin muncul, seperti Hitam, Merah, Coklat, Putih dan Abu-abu. Kemudian, berikan kode untuk masing-masing kategori: 1 untuk "Hitam," 2 untuk "Merah," dan seterusnya. Setelah mengumpulkan data, Kita dapat memasukkan informasi warna rambut setiap responden ke dalam kategori yang sesuai, seperti "1" untuk responden yang berwarna rambut hitam, "2" untuk responden yang berwarna rambut merah, dan seterusnya.

Dengan mengikuti langkah-langkah di atas, Kita dapat mengidentifikasi dan menggolongkan data ke dalam kategori pada skala nominal dengan benar. Hal ini akan membantu dalam melakukan analisis yang tepat dan memberikan informasi yang relevan tentang variabel dalam penelitian kita (Qomari, 1970).

2.2.4. Teknik analisis yang sesuai untuk data skala nominal

Salah satu teknik analisis yang umum digunakan untuk data skala nominal adalah uji chi-square atau uji chi-kuadrat. Uji chi-square digunakan untuk menguji hubungan atau asosiasi antara dua variabel kategorikal atau untuk menguji apakah distribusi data kategorikal berbeda secara signifikan (Wibowo, 2017).

Langkah-langkah umum dalam melakukan uji chi-square untuk data skala nominal:

1. Menyiapkan Data.
   Pastikan data telah diidentifikasi dan dikategorikan dengan benar sesuai dengan skala nominal.
2. Menyusun Tabel Kontingensi.
   Susun data dalam bentuk tabel kontingensi dua arah (2x2 atau lebih) yang menunjukkan frekuensi atau jumlah observasi di setiap sel atau sel-sel data untuk setiap kombinasi kategori dari dua variabel kategorikal yang diuji.
3. Menghitung Statistik Uji.
   Hitung nilai statistik chi-square (χ^2) dengan menggunakan formula yang sesuai berdasarkan tabel kontingensi yang disusun.
4. Menentukan Derajat Kebebasan (df).
   Tentukan derajat kebebasan (df) untuk statistik chi-square berdasarkan ukuran tabel kontingensi. Derajat kebebasan digunakan dalam menghitung nilai probabilitas atau p-value.
5. Tentukan Nilai Signifikansi (p-value).
   Gunakan nilai statistik chi-square dan derajat kebebasan untuk menentukan nilai signifikansi (p-value) dari uji chi-

square. P-value mengindikasikan probabilitas terjadinya hasil yang sama ekstrem atau lebih ekstrem daripada yang diamati, jika hipotesis nol benar.

6. Interpretasi Hasil.
   a. Jika nilai p-value lebih kecil dari tingkat signifikansi yang ditentukan sebelumnya (biasanya 0.05), maka kita dapat menolak hipotesis nol dan menyimpulkan bahwa ada hubungan atau asosiasi yang signifikan antara kedua variabel kategorikal.
   b. Jika nilai p-value lebih besar dari tingkat signifikansi yang ditentukan, maka kita gagal menolak hipotesis nol dan menyimpulkan bahwa tidak ada hubungan yang signifikan antara kedua variabel kategorikal.

Uji chi-square dapat digunakan untuk menguji berbagai hipotesis dan pertanyaan penelitian yang melibatkan variabel kategorikal. Misalnya, menguji apakah ada perbedaan signifikan dalam preferensi politik antara kelompok etnis yang berbeda, apakah ada hubungan antara pendidikan dan preferensi pemilihan partai politik atau apakah ada asosiasi antara jenis kelamin dan preferensi produk tertentu.

2.3. Skala Ordinal

Skala ordinal adalah salah satu jenis skala pengukuran dalam statistika yang digunakan untuk menggolongkan data ke dalam kelompok-kelompok atau kategori-kategori berdasarkan tingkatan atau peringkat. Data pada skala ordinal memiliki urutan yang dapat diurutkan, tetapi tidak memiliki jarak atau selisih yang terukur antara nilai-nilai data (Pentury, Aulele and Wattimena, 2016).

Ciri utama dari skala ordinal adalah bahwa kita hanya dapat mengetahui perbandingan kualitatif antara kategori-kategori atau peringkat, tetapi tidak dapat menentukan seberapa besar perbedaan antara nilai-nilai tersebut secara kuantitatif.

Contoh skala ordinal meliputi:

1. Tingkat Pendidikan.

Misalnya kategori : Tidak Tamat SD, SD, SMP, SMA, Diploma, Sarjana dan seterusnya.

2. Tingkat Kepuasan.
   Misalnya kategori : Sangat Tidak Puas, Tidak Puas, Cukup Puas, Puas dan Sangat Puas.
3. Peringkat Prestasi.
   Misalnya peringkat : pertama, kedua, ketiga dan seterusnya dalam sebuah kompetisi.

Dalam analisis data, untuk data yang diukur dalam skala ordinal, kita dapat melakukan analisis yang melibatkan urutan peringkat, seperti mencari peringkat teratas atau peringkat terbawah. Kita juga dapat menggunakan beberapa teknik statistik yang sesuai, seperti uji nonparametrik seperti uji Mann-Whitney atau uji Wilcoxon untuk membandingkan dua kelompok atau uji Kruskal-Wallis untuk membandingkan lebih dari dua kelompok pada variabel ordinal (Sriwidadi, 2011). Perlu diingat bahwa kita tidak dapat melakukan operasi matematika seperti penjumlahan atau pengurangan pada data dalam skala ordinal karena tidak ada informasi tentang selisih antara nilai-nilai tersebut. Skala ordinal lebih cocok digunakan untuk analisis perbandingan kualitatif atau untuk menggolongkan data berdasarkan tingkatan peringkat.

2.3.1. Perbedaan skala nominal dan ordinal

Perbedaan antara skala nominal dan ordinal terletak pada tingkat informasi yang dapat diberikan oleh masing-masing skala terhadap data yang diukur. Beberapa perbedaan utama antara skala nominal dan ordinal:

1. **Urutan Nilai.**
   a. Skala Nominal.
      Pada skala nominal, data dikelompokkan ke dalam kategori atau kelompok-kelompok yang bersifat saling eksklusif. Data tidak memiliki urutan atau tingkatan yang terukur antara kategori-kategori tersebut.
   b. Skala Ordinal.

Pada skala ordinal, data memiliki urutan atau tingkatan peringkat yang dapat diurutkan. Data memiliki hubungan urutan, tetapi tidak memberikan informasi tentang seberapa besar perbedaan antara nilai-nilai data.

2. **Operasi Matematika**.
   a. Skala Nominal.
      Tidak ada operasi matematika yang dapat dilakukan pada data dalam skala nominal karena kategori-kategori tersebut tidak dapat dijumlahkan, dikurangkan atau dioperasikan dengan cara matematika lainnya.
   b. Skala Ordinal.
      Meskipun data pada skala ordinal memiliki urutan, operasi matematika seperti penjumlahan dan pengurangan tidak berlaku. Kita hanya dapat melakukan operasi yang berkaitan dengan perbandingan kualitatif atau ordinal, seperti menentukan peringkat tertinggi atau peringkat terendah.
3. **Tingkat Informasi**.
   a. Skala Nominal.
      Skala nominal memberikan informasi tentang kategori atau kelompok-kelompok data yang ada dan membantu mengidentifikasi perbedaan kualitatif antara kategori tersebut. Skala nominal tidak memberikan informasi tentang perbandingan atau tingkatan yang bermakna antara kategori-kategori tersebut.
   b. Skala Ordinal.
      Skala ordinal memberikan informasi tentang urutan atau peringkat data yang diukur. Kita dapat mengetahui perbandingan relatif antara kategori-kategori atau peringkat, tetapi tidak dapat menentukan perbedaan yang tepat antara nilai-nilai data.
4. **Contoh.**
   a. Skala Nominal.
      - Variabel jenis kelamin : Laki-laki/Perempuan,
      - Status perkawinan : Belum Menikah, Menikah, Cerai, Duda, Janda.
   b. Skala Ordinal.

- Variabel tingkat pendidikan : Tidak Tamat SD, SD, SMP, SMA, Diploma, Sarjana.
- Tingkat kepuasan : Sangat Tidak Puas, Tidak Puas, Cukup Puas, Puas, Sangat Puas.

Dalam pemilihan metode analisis dan interpretasi data, penting untuk memahami perbedaan antara skala nominal dan ordinal. Penggunaan metode analisis yang sesuai akan membantu memahami informasi yang tepat dari data dan mencegah kesalahan interpretasi.

2.3.2. Urutan dan Peringkat dalam skala Ordinal

Pengertian tentang urutan dan peringkat berkaitan dengan sifat data yang memiliki tingkatan atau derajat perbandingan yang dapat diurutkan. Data pada skala ordinal memiliki nilai-nilai yang dapat diurutkan secara berjenjang, tetapi tidak memiliki informasi tentang jarak atau selisih antara nilai-nilai tersebut.

1. Urutan (Order).
   Urutan mengacu pada pengaturan nilai-nilai data dalam urutan peringkat. Kita dapat mengidentifikasi peringkat tertinggi, peringkat terendah dan peringkat di antara nilai-nilai data. Misalnya 'Tingkat Pendidikan,' dapat diurutkan dari tingkat pendidikan terendah misalnya 'Tidak Tamat SD' hingga tingkat pendidikan tertinggi, misalnya 'Sarjana'.
2. Peringkat (Ranking).
   Peringkat merujuk pada posisi relatif atau tingkatan nilai-nilai data dalam urutan. Data memberikan informasi tentang peringkat atau perbandingan kualitatif antara nilai-nilai tersebut. Misalnya, 'Tingkat Kepuasan', kita dapat menentukan apakah responden memberikan tingkat kepuasan yang berbeda, seperti Sangat Tidak Puas, Tidak Puas, Cukup Puas, Puas atau Sangat Puas.

Dengan adanya urutan dan peringkat data skala ordinal, kita dapat memahami relasi perbandingan kualitatif antara kategori - kategori data. Misalnya, kita dapat mengetahui bahwa tingkat kepuasan "Sangat Puas" lebih tinggi daripada tingkat kepuasan

"Cukup Puas," tetapi tidak menyediakan informasi tentang seberapa besar perbedaan antara tingkat kepuasan tersebut (Ukkas, 2017).

2.3.3. Statistik yang tepat untuk skala ordinal

Penggunaan statistik yang tepat adalah metode analisis yang sesuai dengan sifat data. Karena data dalam skala ordinal memiliki tingkatan peringkat yang dapat diurutkan, tetapi tidak memiliki jarak antara nilai-nilai data, teknik analisis yang paling sesuai adalah metode analisis nonparametrik (Sukarsa and Srinadi, 2012).

Beberapa statistik dan teknik analisis yang tepat untuk data skala ordinal:

1. **Uji Mann-Whitney (Wilcoxon Rank-Sum Test)**.
   Uji Mann-Whitney digunakan untuk membandingkan dua kelompok independen pada variabel ordinal. Uji ini menguji apakah peringkat dari dua kelompok berbeda secara signifikan.
2. **Uji Wilcoxon Signed-Rank.**
   Uji Wilcoxon digunakan untuk membandingkan dua pengukuran terkait pada variabel ordinal dalam kelompok yang sama. Uji ini menguji apakah ada perbedaan yang signifikan antara pengukuran yang berpasangan.
3. **Uji Kruskal-Wallis.**
   Uji Kruskal-Wallis adalah uji nonparametrik yang digunakan untuk membandingkan lebih dari dua kelompok independen pada variabel ordinal. Uji ini menentukan apakah ada perbedaan yang signifikan di antara peringkat kelompok-kelompok tersebut.
4. **Uji Friedman**.
   Uji Friedman digunakan untuk membandingkan lebih dari dua pengukuran terkait pada variabel ordinal dalam kelompok yang sama. Uji ini menguji apakah ada perbedaan yang signifikan di antara peringkat pengukuran yang berpasangan.
5. **Koefisien Korelasi Spearman**.

Koefisien korelasi Spearman digunakan untuk mengukur hubungan antara dua variabel ordinal. Koefisien ini mengukur tingkat asosiasi peringkat antara dua variabel.

Penggunaan statistik nonparametrik pada data skala ordinal merupakan pilihan yang tepat karena metode nonparametrik tidak bergantung pada asumsi distribusi tertentu pada data. Selain itu, metode nonparametrik lebih tahan terhadap data yang tidak terdistribusi normal dan memanfaatkan informasi yang tersedia pada tingkatan peringkat data, meskipun tidak memiliki informasi tentang jarak antara nilai-nilai tersebut.

2.4. Skala Interval

Skala interval adalah salah satu jenis skala pengukuran dalam statistika yang memiliki tingkatan peringkat, jarak yang terukur antara nilai-nilai data dan titik nol yang bukan bersifat mutlak. Perbedaan antara dua nilai adalah konstan dan dapat diukur, tetapi tidak ada nilai nol yang benar-benar menunjukkan ketiadaan dari variabel yang diukur (handayani, Tri; Handayani, Ita; Ikasari, 2019). Ciri utama skala interval adalah bahwa kita dapat melakukan operasi matematika seperti penjumlahan dan pengurangan pada data, tetapi tidak dapat melakukan operasi perkalian atau pembagian karena tidak ada titik nol yang mutlak. Skala interval memungkinkan untuk menentukan perbedaan antara dua nilai dalam satuan yang sama, tetapi tidak memungkinkan untuk menentukan rasio atau perbandingan antara nilai-nilai tersebut.

Contoh skala interval meliputi:

1. **Skala Temperatur Celsius**.
   Pada skala ini perbedaan antara dua angka adalah konstan. Misal 1 derajat Celsius, tetapi tidak ada nilai nol yang benar-benar menunjukkan ketiadaan suhu.
2. **Tahun Kalender (Masehi)**.
   Pada skala tahun kalender, perbedaan antara dua tahun adalah tetap, tetapi tidak ada titik nol yang menunjukkan ketiadaan tahun.

Dalam analisis data, kita dapat melakukan berbagai operasi matematika seperti penjumlahan, pengurangan dan perhitungan rata-rata. Selain itu, juga dapat menggunakan statistik deskriptif seperti rata-rata, standar deviasi dan korelasi Pearson untuk menganalisis data dalam skala interval (Wijayanti *et al.*, 2022). Skala interval memiliki tingkatan peringkat dan jarak yang terukur, tetapi tidak ada nilai nol yang mutlak. Oleh karena itu, perlu berhati-hati dalam menginterpretasikan hasil analisis dan tidak boleh membuat kesimpulan yang tidak tepat tentang perbandingan rasio antara nilai-nilai data.

2.4.1. Sifat khusus skala interval

Sifat khusus data pada skala interval adalah memiliki tingkatan peringkat, jarak yang terukur antara nilai-nilai data dan adanya titik nol yang bukan bersifat mutlak. Sifat khusus skala interval antara lain sebagai berikut:

1. **Tingkatan Peringkat**.
   Data dapat diurutkan secara berjenjang atau berperingkat. Artinya kita dapat mengidentifikasi peringkat atau urutan relatif dari nilai-nilai data, dari yang terkecil hingga yang terbesar. Misalnya Suhu dalam Celsius, dapat diurutkan dari suhu terendah hingga suhu tertinggi.
2. **Jarak yang Terukur.**
   Pada skala interval perbedaan antara dua nilai adalah konstan dan dapat diukur dengan satuan yang sama. Misalnya "Skala Suhu Celsius," perbedaan antara 10°C dan 20°C adalah sama dengan perbedaan antara 20°C dan 30°C yaitu 10 derajat Celsius.
3. **Titik Nol yang Bukan Bersifat Mutlak.**
   Menandakan bahwa titik nol pada skala interval bukan merupakan titik yang benar-benar menunjukkan ketiadaan dari variabel yang diukur. Titik nol hanya merupakan nilai referensi atau nilai awal dari skala tersebut, bukan merupakan titik nol mutlak seperti yang ditemukan pada skala rasio.

Sifat khusus data pada skala interval memungkinkan kita untuk melakukan berbagai operasi matematika, seperti penjumlahan, pengurangan dan perhitungan rata-rata. Namun, tidak mengizinkan operasi perkalian atau pembagian, karena tidak ada nilai nol mutlak yang dapat digunakan sebagai dasar perbandingan rasio antara nilai-nilai data.

2.4.2. Pengenalan terhadap titik nol pada skala interval

Pengenalan terhadap titik nol dalam skala interval penting untuk memahami interpretasi data yang dinyatakan dalam skala ini. Titik nol bukanlah titik nol mutlak, tetapi merupakan nilai referensi atau nilai awal dari skala. Beberapa penjelasan tentang titik nol dalam skala interval dan bagaimana memahami interpretasi data dalam skala ini:

1. Titik Nol dalam Skala Interval.
   Titik nol hanya merupakan titik acuan atau titik awal dari skala, bukan nilai yang menyatakan ketiadaan dari variabel yang diukur. Misalnya suhu dalam Celsius, suhu 0°C adalah nilai referensi, bukan berarti tidak ada suhu sama sekali. Jika suhu diukur dalam skala Kelvin (yang merupakan skala rasio dengan titik nol mutlak), suhu 0K menunjukkan ketiadaan panas (nol mutlak) atau suhu yang paling rendah.
2. Interpretasi Data dalam Skala Interval:
   a. Ketika memahami data dalam skala interval, kita dapat mengurutkan nilai-nilai data berdasarkan tingkatan peringkat dan mengukur jarak antara nilai-nilai tersebut. Misalnya, jika kita memiliki data suhu dalam skala Celsius (°C), kita dapat mengurutkan suhu dari yang terendah hingga yang tertinggi dan kita dapat mengukur selisih suhu antara dua titik.
   b. Dalam skala interval, kita tidak dapat melakukan perbandingan rasio atau menyatakan bahwa suatu nilai dua kali lebih besar daripada nilai lainnya. Misalnya, tidak tepat untuk mengatakan bahwa suhu 40°C dua kali

lebih panas daripada suhu 20°C karena skala interval tidak memiliki titik nol mutlak yang menunjukkan ketiadaan suhu.

c. Oleh karena itu, saat menginterpretasikan data dalam skala interval, hindari penggunaan operasi perkalian, pembagian atau perbandingan rasio, karena data dalam skala interval tidak mengizinkan operasi tersebut. Sebagai gantinya, fokuskan pada perbandingan kualitatif atau peringkat antara nilai-nilai data dan gunakan metode analisis yang sesuai, seperti uji t-Student untuk membandingkan dua kelompok independen atau analisis regresi untuk menilai hubungan antara variabel-variabel interval.

2.4.3. Penggunaan Statistik yang tepat pada skala interval

Statistik yang tepat digunakan adalah statistik parametrik. Statistik parametrik adalah metode analisis yang memerlukan asumsi tentang distribusi data dan memanfaatkan informasi tentang jarak antara nilai-nilai data. Data Skala interval memiliki tingkatan peringkat dan jarak yang terukur, sehingga statistik parametrik dapat memberikan analisis yang lebih kuat dan lebih tepat jika asumsi yang sesuai dipenuhi.

Beberapa teknik analisis statistik yang tepat untuk data skala interval:

1. **Uji t-Student**.
   Uji t-Student digunakan untuk membandingkan dua kelompok independen pada variabel interval. Kita dapat menggunakan uji t-Student untuk membandingkan rata-rata nilai ujian antara siswa laki-laki dan perempuan.
2. **Uji Anova (Analysis of Variance)**.
   Uji Anova digunakan untuk membandingkan lebih dari dua kelompok independen pada variabel interval. Kita dapat menggunakan uji Anova untuk membandingkan rata-rata pendapatan antara tiga kelompok pekerja berbeda.
3. **Regresi Linier**

Analisis ini digunakan untuk menilai hubungan linier antara dua variabel interval. Misal menggunakan analisis regresi linier untuk memahami bagaimana penjualan produk dipengaruhi oleh biaya iklan.

4. **Korelasi Pearson**.
   Korelasi Pearson digunakan untuk mengukur kekuatan dan arah hubungan linier antara dua variabel interval. Korelasi Pearson memberikan koefisien korelasi yang berkisar antara -1 hingga +1, dengan nilai positif menunjukkan korelasi positif, nilai negatif menunjukkan korelasi negatif dan nilai mendekati nol menunjukkan korelasi lemah.
5. **Uji Regresi Berganda**.
   Uji regresi berganda adalah perluasan dari analisis regresi linier yang memungkinkan untuk mengevaluasi hubungan antara satu variabel dependen dan dua atau lebih variabel independen yang berhubungan dengan variabel dependen.

Penting untuk memastikan bahwa data Kita memenuhi asumsi statistik parametrik sebelum menerapkan teknik analisis di atas. Asumsi tersebut termasuk asumsi normalitas, homogenitas varians dan independensi data. Jika asumsi tidak terpenuhi, mungkin perlu menggunakan metode analisis nonparametrik yang lebih sesuai (Jaya and Ardat, 2013).

2.5. Skala Rasio

Skala rasio adalah salah satu jenis skala pengukuran dalam statistik mempunyai pengukuran yang paling lengkap dan informatif. Skala rasio memiliki sifat yang unik karena memiliki nol mutlak dan memungkinkan operasi aritmatika yang lebih kompleks. Pada skala rasio, semua jenis operasi matematika dapat dilakukan, termasuk penjumlahan, pengurangan, perkalian dan pembagian, hasilnya akan memiliki interpretasi yang jelas dan relevan (David, Bakrie and Bakrie, 2018).

Ciri-ciri skala rasio:

1. Memiliki titik nol (nol mutlak).

Skala rasio memiliki titik nol yang mutlak yaitu nol pada skala ini bukan sekadar nilai relatif atau tidak ada nilainya, akan tetapi merupakan nilai mutlak yang benar-benar menyatakan ketiadaan dari atribut yang diukur. Perbandingan rasio antara dua nilai pada skala ini memiliki arti yang bermakna.

2. Memiliki interval yang jelas.
   Selain memiliki titik nol, skala rasio juga memiliki interval yang jelas antara nilai-nilainya. Interval ini menyatakan perbedaan antara dua nilai pada skala dan dapat diukur secara konsisten.

Contoh skala rasio:

1. **Berat Badan**.
   Nilai nol dalam skala ini menunjukkan tidak ada berat badan dan perbandingan rasio antara dua nilai berat badan memiliki arti yang jelas. Misalnya, jika berat badan seseorang adalah 60 kg dan berat badan orang lain adalah 30 kg, maka berat badan orang pertama adalah dua kali berat badan orang kedua.
2. **Jarak.**
   Skala rasio juga digunakan untuk mengukur jarak. Nilai nol dalam skala ini menunjukkan titik awal atau titik acuan dan perbandingan rasio antara dua jarak memiliki arti yang relevan. Misal, jarak dua kota pertama adalah 100 km dan jarak antara kota kedua adalah 50 km, maka jarak antara dua kota pertama dua kali lebih besar dari jarak antara kota kedua.
3. **Waktu.**
   Skala rasio juga digunakan untuk mengukur waktu. Nilai nol dalam skala waktu menunjukkan titik awal atau awal hitungan waktu dan perbandingan rasio antara dua nilai waktu memiliki arti yang jelas. Misal, waktu perjalanan dari A ke B adalah 4 jam dan waktu perjalanan dari B ke C adalah 2 jam, maka waktu perjalanan dari A ke B dua kali lebih lama daripada waktu perjalanan dari B ke C.

Skala rasio memberikan fleksibilitas dalam analisis data dan memungkinkan untuk mengambil kesimpulan yang lebih mendalam dan akurat. Namun, penting untuk memahami jenis skala yang digunakan dalam pengukuran data, karena jenis skala akan mempengaruhi jenis analisis dan interpretasi yang dapat dilakukan terhadap data.

2.5.1. Karakteristik data pada skala rasio.

Karakteristik data pada skala rasio memiliki beberapa ciri khas yang membedakannya dari jenis skala pengukuran lainnya. Berikut adalah beberapa karakteristik utama dari data pada skala rasio:

1. Titik Nol yang Mutlak.
   Salah satu ciri utama dari skala rasio adalah adanya titik nol yang mutlak. Titik nol ini merupakan titik acuan yang sebenarnya dari atribut yang diukur. Misalnya, pada skala rasio berat badan, nilai nol (0 kg) menunjukkan bahwa tidak ada berat badan atau berat badan sama dengan nol.

2. Properti Perbandingan yang Bermakna.
   Perbandingan antara dua nilai memiliki arti yang jelas dan relevan. Jika nilai A adalah dua kali lebih besar dari nilai B, maka ini berarti nilai A memiliki dua kali lipat dari jumlah atribut yang diukur dibandingkan dengan nilai B. Properti perbandingan yang bermakna ini memungkinkan operasi matematika seperti perkalian dan pembagian untuk dilakukan dengan interpretasi yang konsisten.

3. Interval yang Konsisten.
   Interval ini mengindikasikan perbedaan atau jarak yang sama antara dua nilai pada skala. Misal, perbedaan antara 20 dan 30 memiliki interval yang sama dengan perbedaan antara 70 dan 80. Interval yang konsisten memungkinkan perhitungan aritmatika yang valid dan konsisten.

4. Pengukuran Absolut.
   Artinya, pada skala ini, dapat diukur besaran atau perbandingan dari atribut yang diukur secara nyata.

Contohnya, jika berat badan seseorang adalah 40 kg, berat badan tersebut merupakan angka absolut yang menunjukkan bobot fisik yang sesungguhnya.

Data pada skala rasio memberikan tingkat informasi yang paling lengkap dan mendalam karena memiliki titik nol yang mutlak dan memungkinkan perbandingan yang bermakna antara nilai-nilainya. Sehingga analisis dan interpretasi data pada skala rasio dapat memberikan pemahaman yang lebih akurat dan mendalam terhadap atribut yang diukur.

2.5.2. Analisis data pada skala rasio.

Analisis data pada skala rasio memungkinkan kita untuk menggunakan berbagai metode statistik yang lebih canggih dan memberikan pemahaman yang lebih mendalam tentang hubungan antara variabel (wulansari, 2016). Berikut langkah-langkah untuk melakukan analisis data pada skala rasio dengan tepat:

1. **Deskripsi Data**.
   Deskripsi data yaitu menyusun ringkasan statistik untuk setiap variabel pada skala rasio. Ringkasan statistik ini meliputi nilai rata-rata (mean), median, simpangan baku (standard deviation), minimum dan maksimum. Dengan melihat ringkasan statistik ini, kita dapat memahami distribusi data dan mengidentifikasi nilai-nilai ekstrim atau outlier yang mungkin mempengaruhi hasil analisis.
2. **Visualisasi Data**.
   Data pada skala rasio dapat divisualisasikan menggunakan grafik atau diagram yang sesuai. Misalnya, untuk menganalisis distribusi data, histogram atau box plot dapat digunakan. Jika ingin melihat hubungan antara dua variabel, scatter plot atau line plot dapat digunakan. Visualisasi data membantu kita memahami pola dan tren dalam data, serta mengidentifikasi anomali atau pola yang menarik.
3. **Uji Hipotesis**.

Pada analisis data pada skala rasio, kita dapat menggunakan uji hipotesis untuk menguji perbedaan antara dua atau lebih kelompok data. Uji t, uji ANOVA (Analysis of Variance) dan uji-regresi adalah contoh dari beberapa uji hipotesis yang dapat digunakan untuk analisis data pada skala rasio. Uji hipotesis membantu kita menentukan apakah perbedaan yang kita amati antara kelompok-kelompok data adalah signifikan secara statistik atau hanya hasil kebetulan.

4. **Korelasi dan Regresi.**
   Analisis korelasi dan regresi sangat relevan dalam analisis data pada skala rasio. Korelasi digunakan untuk mengukur hubungan linear antara dua variabel, sementara regresi digunakan untuk memodelkan hubungan kausal antara variabel dependen dan independen. Analisis korelasi dan regresi membantu kita memahami sejauh mana satu variabel mempengaruhi variabel lain dan memberikan gambaran tentang kekuatan dan arah hubungan antara variabel-variabel tersebut.

5. **Analisis Multivariat.**
   Jika terdapat banyak variabel pada data skala rasio, analisis multivariat, seperti analisis faktor atau analisis klaster, dapat digunakan untuk mengidentifikasi pola dan struktur yang kompleks dalam data. Analisis ini membantu kita menyederhanakan data dan mengelompokkan variabel-variabel yang saling terkait dalam faktor-faktor atau kelompok-kelompok yang lebih bermakna.

6. **Interpretasi Hasil.**
   Setelah melakukan analisis data, langkah terakhir adalah menginterpretasi hasil analisis dengan hati-hati dan kritis. Hasil analisis harus dihubungkan kembali ke pertanyaan penelitian atau tujuan analisis awal. Interpretasi yang tepat akan memberikan wawasan yang berarti dan dapat digunakan sebagai dasar untuk pengambilan keputusan atau tindakan lebih lanjut.

Dalam melakukan analisis data pada skala rasio, penting untuk menggunakan metode statistik yang sesuai dan menghindari kesalahan dalam interpretasi hasil (Nurizzati, 2017).

2.6. Transfomasi Data

Transformasi data adalah proses mengubah atau mengolah data secara matematis agar data lebih memenuhi asumsi statistik tertentu, atau agar analisis statistik yang dilakukan memberikan hasil yang lebih akurat dan bermakna (Sappaile and Makassar, 2007).

Alasan mengapa transformasi data mungkin diperlukan dalam analisis statistik adalah sebagai berikut:

1. **Mengatasi Ketidaknormalan Distribusi.**
   Beberapa metode statistik seperti uji parametrik seperti uji t dan analisis varians (ANOVA), mengasumsikan distribusi data normal. Jika data tidak memenuhi asumsi distribusi normal, transformasi data dapat digunakan untuk mendekati atau menciptakan distribusi yang lebih mendekati normal.
2. **Stabilisasi Variansi**.
   Transformasi data juga digunakan untuk mengatasi masalah heteroskedastisitas yaitu ketika variansi dari data tidak konstan di seluruh rentang nilai. Transformasi data dapat mengstabilkan variansi dan membuat data lebih homogen.
3. **Linearity**.
   Beberapa analisis statistik seperti analisis regresi linear, mengasumsikan hubungan linear antara variabel dependen dan independen. Jika hubungan tersebut tidak linear, transformasi data dapat membantu mengubahnya menjadi hubungan yang lebih linier.
4. **Reduksi Skala**.
   Dalam beberapa kasus, data memiliki skala yang sangat besar atau bervariasi, sehingga analisis statistik mungkin sulit diinterpretasikan. Transformasi data dapat mengurangi skala data dan membuatnya lebih mudah untuk dipahami dan dibandingkan.
5. **Normalisasi Data.**

Transformasi data juga dapat digunakan untuk mengubah data ke dalam rentang nilai yang lebih terbatas atau normalisasi. Ini berguna dalam situasi di mana variabel yang berbeda memiliki unit atau skala yang berbeda dan perlu dibandingkan secara relatif.

2.6.1. Metode Transformasi.

Beberapa metode transformasi data yang umum digunakan adalah sebagai berikut:

1. **Transformasi Logaritma**.
   Metode ini digunakan untuk mengatasi ketidaknormalan distribusi data yang sangat miring ke kanan atau memiliki variansi yang tidak stabil. Mengunakan logaritma data dapat meratakan distribusi dan mengurangi variansi data. Transformasi logaritma umumnya digunakan pada data yang bernilai positif, seperti harga, pendapatan dan ukuran populasi.
2. **Transformasi Aritmatika.**
   Transformasi ini mencakup akar kuadrat, akar pangkat n (misalnya akar pangkat 2 atau 3) atau transformasi lainnya yang bersifat aritmatika. Metode ini sering digunakan untuk mengurangi variansi data dan mengatasi masalah ketidaknormalan distribusi.
3. **Transformasi Standard.**
   Metode ini digunakan untuk mengubah data menjadi z-skor (standard skor) dengan rata-rata nol dan simpangan baku satu. Transformasi ini membantu dalam membandingkan data secara relatif dan mengatasi masalah skala yang berbeda pada variabel.
4. **Transformasi Invers.**
   Digunakan untuk mengubah data yang memiliki hubungan invers menjadi hubungan linear. Misalnya, jika data memiliki hubungan berbanding terbalik, transformasi invers dapat membantu mengubahnya menjadi hubungan linier.
5. **Transformasi Kategorikal.**

Jika data pada variabel adalah kategorikal atau nominal, transformasi dapat dilakukan dengan mengubah variabel menjadi variabel dummy. Variabel dummy adalah variabel biner yang bernilai 0 atau 1, yang digunakan untuk mewakili kategori atau kelompok tertentu dalam analisis.

6. **Normalisasi.**
   Metode normalisasi digunakan untuk mengubah data ke dalam rentang nilai tertentu, biasanya dari 0 hingga 1. Normalisasi berguna jika data pada variabel memiliki skala yang sangat berbeda atau memiliki satuan yang berbeda.

Pemilihan metode transformasi data harus didasarkan pada pemahaman yang baik tentang asumsi statistik, tujuan analisis dan karakteristik data. Sebelum melakukan transformasi data, terlebih dahulu pahami tujuan transformasi dan dampaknya pada interpretasi hasil analisis. Selain itu, perlu diingat bahwa transformasi data dapat mengubah distribusi atau skala data, sehingga interpretasi hasil setelah transformasi harus dilakukan dengan hati-hati.

2.7. Pemilihan Skala yang Tepat

Memilih skala yang tepat dalam mengumpulkan data untuk penelitian ilmu sosial sangat penting untuk memastikan bahwa data yang dikumpulkan berkualitas dan sesuai dengan pertanyaan penelitian yang ingin dijawab (Nana and Elin, 2018). Berikut adalah beberapa pedoman untuk memilih skala yang tepat:

1. **Tujuan Penelitian**.
   Pertama-tama, identifikasi tujuan penelitian dengan jelas. Pahami pertanyaan penelitian yang ingin di jawab dan jenis informasi yang perlu dikumpulkan. Pertimbangkan apakah Kita ingin mengukur frekuensi, intensitas, peringkat, preferensi atau hubungan antara variabel.
2. **Jenis Variabel.**
   Tentukan jenis variabel yang ingin diamati. Variabel dapat berupa kategori (nominal), urutan (ordinal), interval atau rasio. Pemilihan skala harus sesuai dengan jenis variabel

yang ingin diukur. Misal untuk variabel kategori, skala nominal dapat digunakan, sementara untuk variabel urutan, skala ordinal lebih cocok.

3. **Tingkat Pengukuran.**
   Pertimbangkan tingkat pengukuran data yang dibutuhkan. Tingkat pengukuran mengacu pada tingkat presisi atau kemampuan skala untuk memberikan informasi yang relevan. Tingkat pengukuran dapat berupa nominal, ordinal, interval atau rasio. Pastikan skala yang dipilih sesuai dengan tingkat pengukuran yang diperlukan untuk menjawab pertanyaan penelitian.
4. **Jumlah Opsi pada Skala.**
   Pertimbangkan jumlah opsi atau kategori pada skala. Jika terlalu banyak opsi, responden dapat kesulitan memilih jawaban yang sesuai. Sebaliknya, jika terlalu sedikit opsi, Kita mungkin kehilangan keragaman dalam tanggapan. Pilih jumlah opsi yang cukup untuk mencakup variasi yang relevan, namun tidak terlalu banyak hingga menyulitkan responden.
5. **Eksplorasi dan Uji Coba**.
   Sebelum mengumpulkan data sebenarnya, lakukan eksplorasi dan uji coba skala yang dipilih. Pastikan skala mudah dipahami oleh responden dan mampu menghasilkan data yang relevan dengan pertanyaan penelitian.
6. **Pertimbangkan Konteks Budaya dan Sosial**.
   Jika penelitian melibatkan responden dari budaya atau latar belakang sosial yang berbeda, pertimbangkan konteks budaya dan sosial dalam memilih skala. Pastikan skala yang dipilih dapat dipahami dan relevan bagi responden yang beragam.
7. **Pertimbangkan Skala Skor dan Metode Analisis.**
   Pilih skala skor yang sesuai dengan metode analisis yang akan digunakan. Beberapa metode analisis membutuhkan skala rasio, sementara yang lain dapat menggunakan skala ordinal atau interval. Pastikan skala yang dipilih sesuai dengan metode analisis yang ingin digunakan.

Dengan memperhatikan pedoman di atas, Kita dapat memilih skala yang tepat untuk mengumpulkan data dalam penelitian ilmu sosial. Pastikan bahwa data yang dikumpulkan akurat, relevan dan dapat memberikan wawasan yang berarti untuk menjawab pertanyaan penelitian.

2.7.1. Etika dan metodologi dalam pemilihan skala pengukuran.

Pertimbangan etika dan metodologi dalam pemilihan skala pengukuran merupakan hal penting dalam penelitian untuk memastikan integritas, validitas, dan kehandalan data yang dikumpulkan. Beberapa pertimbangan penting dalam hal ini:

Pertimbangan Etika.

1. Perlindungan Privasi dan Keamanan Data.
   Pastikan bahwa data yang dikumpulkan menghormati privasi responden dan data disimpan dengan aman. Jaga kerahasiaan informasi pribadi dan pastikan bahwa data tidak dapat diakses oleh pihak yang tidak berwenang.
2. *Informed Consent*.
   Pastikan bahwa semua responden memberikan informed consent sebelum ikut dalam penelitian. Berikan informasi yang jelas tentang tujuan penelitian, penggunaan data, hak-hak responden dan konsekuensi partisipasi.
3. Penghindaran Bias dan Diskriminasi.
   Pastikan bahwa skala pengukuran yang digunakan tidak mengandung pertanyaan atau pernyataan yang dapat menyebabkan bias atau diskriminasi terhadap kelompok atau individu tertentu.
4. Etika Penggunaan Instrumen.
   Jika Kita menggunakan skala pengukuran yang sudah ada (misalnya alat ukur yang telah diuji validitas dan reliabilitas), pastikan untuk mematuhi hak cipta dan mengakui sumbernya secara tepat.

Pertimbangan Metodologi.

1. Validitas.
   Pastikan bahwa skala pengukuran yang digunakan adalah valid untuk mengukur variabel yang ingin diukur. Validitas mengacu pada sejauh mana skala benar-benar mengukur apa yang diinginkan.
2. Reliabilitas.
   Pastikan bahwa skala pengukuran memiliki reliabilitas yang baik, yaitu skala dapat mengukur variabel secara konsisten dan dapat diandalkan dalam mengukur atribut yang sama pada waktu yang berbeda atau di bawah kondisi yang berbeda.
3. Responsiveness.
   Skala pengukuran harus responsif terhadap perubahan yang relevan dalam variabel yang diukur dan harus mampu mendeteksi perubahan atau perbedaan yang signifikan.
4. Relevansi.
   Pastikan bahwa skala yang dipilih relevan dengan pertanyaan penelitian yang ingin dijawab. Dengan kata lain harus mampu menghasilkan data yang relevan dan bermanfaat untuk analisis penelitian.
5. Konsistensi dengan Metode Analisis.
   Pastikan skala pengukuran yang dipilih konsisten dengan metode analisis yang akan digunakan.
6. Kepraktisan.
   Pertimbangkan kepraktisan dapat merujuk skala harus mudah diimplementasikan dan memungkinkan responden menjawab dengan mudah.

Dengan mempertimbangkan aspek etika dan metodologi, pemilihan skala pengukuran yang tepat akan membantu memastikan bahwa data yang dikumpulkan dapat dipercaya, relevan dan memberikan wawasan yang bermakna dalam penelitian.

DAFTAR PUSTAKA

David, W., Bakrie, U. and Bakrie, U. (2018) *Metode Statistik Untuk Ilmu dan Teknologi Pangan*.

handayani, Tri; Handayani, Ita; Ikasari, I.H. (2019) 'Buku Statistika Dasar', *Angewandte Chemie International Edition, 6(11), 951–952.*, pp. 5–24.

Ilmiyah, T. (2013) '"Pengaruh Pemanfaatan Koleksi Local Content Terhadap Kegiatan Penelitian Mahasiswa Yang Sedang Mengerjakan Skripsi/Tugas Akhir Di Perpustakaan Fakultas Ilmu Budaya Universitas Diponegoro Semarang"', *Ilmu Perpustakaan*, 2(2), p. 6.

Jaya, I. and Ardat (2013) *Penerapan Statistik Untuk Pendidikan*. Available at: http://repository.uinsu.ac.id/2500/1/ISI PENERAPAN STATISTIK ARDAT.pdf.

Junaidi (2015) 'Memahami skala-skala pengukuran', *Research Gate*, (May), pp. 1–5.

McHugh, M.L. (2012) 'The Chi-square test of independence', *Biochemia Medica*, 23(2), pp. 143–149. Available at: https://doi.org/10.11613/BM.2013.018.

Nana, D. and Elin, H. (2018) 'Memilih Metode Penelitian Yang Tepat: Bagi Penelitian Bidang Ilmu Manajemen', *Jurnal Ekonologi Ilmu Manajemen*, 5(1), p. 288. Available at: https://jurnal.unigal.ac.id/index.php/ekonologi/article/view/1359.

Nurizzati, Y. (2017) 'Peranan Statistika Dalam Penelitian Sosial Kuantitatif', *Jurnal SAINTEKOM*, 6(2), pp. 91–105.

Pentury, T., Aulele, S.N. and Wattimena, R. (2016) 'Analisis Regresi Logistik Ordinal', *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 10(1), pp. 55–60. Available at: https://doi.org/10.30598/barekengvol10iss1pp55-60.

Qomari, R. (1970) 'Teknik Penelusuran Analisis Data Kuantitatif dalam Penelitian Kependidikan', *INSANIA : Jurnal Pemikiran Alternatif Kependidikan*, 14(3), pp. 527–539. Available at: https://doi.org/10.24090/insania.v14i3.372.

Sappaile, B.I. and Makassar, U.N. (2007) 'Oleh : Baso Intang Sappaile) penilaian yang sifatnya ordinal , seperti skala Likert . Skor butir pernyataan pada skala ordinal tidaklah tepat dilakukan penjumlahan dari sejumlah skor , tetapi penjumlahan skor dapat dilakukan bila skor pernyataan merupa', *Jurnal Pendidikan Dan Kebudayaan*, 0215–2673(January), pp. 126–135.

Sriwidadi, T. (2011) 'Penggunaan Uji Mann-Whitney pada Analisis Pengaruh Pelatihan Wiraniaga dalam Penjualan Produk Baru', *Binus Business Review*, 2(2), p. 751. Available at: https://doi.org/10.21512/bbr.v2i2.1221.

Sukarsa, I.K.G. and Srinadi, I.G.A.M. (2012) 'Estimator kernel dalam model regresi nonparametrik', *Jurnal Matematika*, 2(1), pp. 19–30.

Ukkas, M.I. (2017) 'Implementasi skala likert pada metode perbandingan eksponensial untuk menentukan pilihan asuransi', *Seminar Nasional Sistem Informasi Indonesia*, (November), p. 101. Available at: http://is.its.ac.id/pubs/oajis/index.php/home/detail/1751/Implementasi-Skala-Likert-Pada-Metode-Perbandingan-Eksponensial-Untuk-Menentukan-Pilihan-Asuransi.

Wibowo, A. (2017) 'Uji Chi-Square pada Statistika dan SPSS', *Jurnal Ilmiah SINUS*, 4(2), p. 38.

Wijayanti, R.R. *et al.* (2022) *Statistik Deskriptif*, *Widina Media Utama*. Available at: www.penerbitwidina.com.

Wulansari, Andhita Dessy. (2016) *Aplikasi Statistika Parametrik dalam Penelitian.* Pustaka Felicha, Yogyakarta. ISBN 978-602-70278-6-2.

BAB 3

UKURAN-UKURAN STATISTIKA

Oleh Marsuddin Musa, S.Si., M.Stat.

3.1. Pendahuluan

Pada BAB sebelumnya telah dibahas mengenai teknik-teknik penyajian data. Masih berkaitan dengan itu, pada BAB ini akan dibahas mengenai statistika deskriptif lainnya, yaitu ukuran-ukuran dalam statistika. Materi ini sangat penting untuk dipelajari karena dalam menganalisis data kuantitatif diperlukan ukuran-ukuran statistika untuk meringkas dan menjelaskan karakteristik dari suatu kumpulan data. Ukuran-ukuran statistika terbagi menjadi ukuran pemusatan dan ukuran penyebaran data. Sebelum membahas ukuran-ukuran tersebut, perlu dipahami terlebih dahulu mengenai notasi matematis yang digunakan untuk mengukur data populasi dan data sampel. Dalam hal ini adalah notasi yang menunjukan suatu parameter dan staistik. Parameter merupakan nilai-nilai yang menjelaskan ciri dari suatu populasi. Sementara statistik yaitu nilai-nilai yang menjelaskan ciri dari suatu sampel (Walpole, 1995).

Beberapa perbedaan notasi antara parameter dan statistik yaitu untuk rata-rata populasi disimbolkan dengan μ dibaca "myu", sementara rata-rata sampel yaitu $\bar{x}$ dibaca "x bar". Standar deviasi dan varians populasi disimbolkan dengan σ dibaca "sigma" dan σ^2, sedangkan standar deviasi dan varians sampel disimbolkan dengan s dan s^2. Selain itu, terdapat notasi matematis yang paling sering digunakan untuk menentukan ukuran-ukuran dalam statistika, yaitu simbol $\sum$ dibaca "sigma" untuk menyatakan penjumlahan, sehingga jika $\sum_{i=1}^{n} x_i$ maka dapat dibaca dengan penjumlahan x_i dimana i berjalan dari 1 sampai n.

3.2. Ukuran Pemusatan Data

Ukuran pemusatan data merupakan nilai tunggal yang menunjukan karakteristik dari data yang mewakili suatu kumpulan data (Suharyadi and Purwanto S.K., 2016). Beberapa ukuran pemusatan data yang paling sering digunakan yaitu rata-rata, median, modus, dan ukuran letak (kuartil, desil, dan persentil). Ukuran-ukuran tersebut telah banyak diterapkan dalam berbagai bidang ilmu, termasuk dalam ilmu-ilmu sosial.

3.2.1. Rata-Rata

Rata-rata merupakan salah satu ukuran yang menunjukkan pusat data yang dapat mewakili karakteristik dari suatu kumpulan data. Rata-rata diperoleh dengan menjumlahkan keseluruhan data dan membaginya dengan banyaknya data (Suharyadi and Purwanto S.K., 2016). Agar mempermudah memahami rata-rata akan dibahas secara berurutan mengenai rata-rata populasi, rata-rata sampel, rata-rata tertimbang, dan rata-rata data berkelompok.

a. **Rata-Rata Populasi**

Rata-rata populasi merupakan nilai rata-rata dari suatu kumpulan data populasi (parameter) yang disimbolkan dengan μ. Berikut adalah rumus yang digunakan untuk menentukan rata-rata populasi.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Keterangan:
μ : rata-rata populasi
$\sum_{i=1}^{N} x_i$: jumlah pengamatan ke-i, dimana i = 1,2,...,N
N : banyaknya data pengamatan dalam populasi.

Contoh 4-1

Sebagai contoh, selama satu bulan terakhir banyaknya penjualan sepeda motor pada lima dealer di Kota A adalah 28, 22, 25, 20, dan 25. Dengan memandang data tersebut adalah populasi hitunglah nilai rata-ratanya.

Penyelesaian:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{28 + 22 + 25 + 20 + 25}{5} = \frac{120}{5} = 24$$

b. **Rata-Rata Sampel**

Rata-rata sampel merupakan nilai rata-rata dari suatu kumpulan data sampel (statistik) yang disimbolkan dengan $\bar{x}$. Berikut adalah rumus yang digunakan untuk menentukan rata-rata sampel.

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

Keterangan:
$\bar{x}$: rata-rata sampel
$\sum_{i=1}^{n} x_i$: jumlah pengamatan ke-*i*, dimana *i* = 1,2,...,*n*
n : banyaknya data pengamatan dalam sampel.

Contoh 4-2

Sebagai contoh, seorang peneliti melakukan pencatatan terhadap bayi yang lahir selama setahun terakhir pada delapan Desa sampel. Hasil pencatatannya adalah 33, 28, 31, 41, 36, 25, 30, dan 32. Hitunglah nilai rata-ratanya.

Penyelesaian:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{33 + 28 + 31 + 41 + 36 + 25 + 30 + 32}{8} = \frac{256}{8} = 32$$

c. **Rata-Rata Tertimbang**

Setiap data pada perhitungan rata-rata populasi dan sampel dianggap mempunyai tingkat atau bobot yang sama. Namun pada beberapa kasus ada data yang dipandang memiliki bobot berbeda. Hal ini karena data dapat mempunyai pengaruh dan kepentingan berbeda, yaitu pengaruh waktu dan pengaruh volume (Suharyadi and Purwanto S.K., 2016). Rata-rata tertimbang dikembangkan sebagai upaya mempertimbang-kan nilai bobot, dan peran dari setiap data yang dianggap

berbeda. Rumus yang digunakan untuk menentukan rata-rata tertimbang adalah sebagai berikut.

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

Keterangan:

$\bar{x}_w$: rata-rata tertimbang
$\sum_{i=1}^{n} w_i$: jumlah bobot pengamatan ke-*i*, dimana *i* = 1,2,...,*n*
$\sum_{i=1}^{n} w_i x_i$: jumlah hasil kali bobot dan nilai pengamatan ke-*i*, dimana *i* = 1,2,...,*n*
n : banyaknya data pengamatan dalam sampel.

Contoh 4-3

Sebagai contoh, CV Makmur Jaya adalah produsen sepatu dengan skala kecil di Kota X. Jumlah produksi dan tenaga kerja selama lima hari kerja adalah:

Tabel 3.1 Data Produksi dan Tenaga Kerja

Jumlah Tenaga Kerja (orang)	Produksi Sepatu (pasang)
5	300
12	1000
7	600
8	900
4	250

(Sumber: Penulis, 2023. Diolah)

Hitunglah rata-rata produksi per hari dengan pembobot (*w*) jumlah tenaga kerja.

Penyelesaian:

$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} = \frac{1.500 + 12.000 + \cdots + 1.000}{5 + 12 + \cdots + 4} = \frac{25.900}{36} = 719$$

d. **Rata-Rata Data Berkelompok**

Rata-rata data berkelompok merupakan nilai rata-rata dari data yang sudah dikelompokkan dalam bentuk distribusi frekuensi. Data yang sudah dikelompokkan akan kehilangan identitas data mentah sehingga untuk melihat nilai rata-rata harus diduga dari distribusi frekuensinya (Suharyadi and Purwanto S.K., 2016). Berikut adalah rumus yang digunakan untuk menentukan rata-rata data berkelompok.

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

Keterangan:

$\bar{x}$: rata-rata data berkelompok
$\sum_{i=1}^{k} f_i$: jumlah frekuensi kelas ke-*i*, dimana *i* = 1,2,...,*k*
$\sum_{i=1}^{k} x_i f_i$: jumlah hasil kali titik tengah dan frekuensi kelas ke-*i*, dimana *i* = 1,2,...,*k*
k : banyaknya kelas.

Contoh 4-4

Sebagai contoh, berikut adalah data net profit yang diperoleh 80 UKM yang ada dikota X selama setahun terakhir.

Tabel 3.2 Data Net Profit UKM

Kelas ke-i	Interval Net Profit (Juta Rp)	Frekuensi (f_i)
1	31 – 40	2
2	41 – 50	3
3	51 – 60	5
4	61 – 70	13
5	71 – 80	24
6	81 – 90	21
7	91 – 100	12

(Sumber: Penulis, 2023. Diolah)

Tentukanlah rata-rata data net profit diatas.

Penyelesaian:

Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.3 Perhitungan Rata-rata Data Berkelompok

Kelas ke-i	Interval Kelas	Titik Tengah (x_i)	f_i	$x_i . f_i$
1	31 – 40	35,5	2	71
2	41 – 50	45,5	3	136,5
3	51 – 60	55,5	5	277,5
4	61 – 70	65,5	13	851,5
5	71 – 80	75,5	24	1812
6	81 – 90	85,5	21	1795,5
7	91 – 100	95,5	12	1146
Jumlah			80	6090

(Sumber: Penulis, 2023. Diolah)

Sehingga diperoleh,

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} = \frac{71 + 136{,}5 + \cdots + 1146}{2 + 3 + \cdots + 12} = \frac{6090}{80} = 76{,}125$$

3.2.2. Median

Median merupakan salah satu ukuran pemusatan data. Median adalah titik tengah dari suatu kumpulan data setelah data diurutkan, baik dari yang terkecil sampai terbesar atau sebaliknya.

a. **Median Data Tidak Berkelompok**

Median untuk data tidak berkelompok adalah nilai yang letaknya ditengah setelah data diurutkan, namun datanya belum dikelompokkan dalam bentuk distribusi frekuensi. Cara menentukan nilai median yaitu:

1. Jika banyaknya pengamatan (n) ganjil, maka nilai median merupakan nilai yang letaknya tepat ditengah setelah data diurutkan. Median = x_j dimana $j = \frac{(n+1)}{2}$; j adalah letak data.

2. Jika banyaknya pengamatan (n) genap, maka nilai median merupakan nilai rata-rata dua pengamatan yang ditengah setelah data diurutkan. Median = $\frac{(x_j+x_{j+1})}{2}$ dimana $j = \frac{n}{2}$

Contoh 4-5

Sebagai contoh, selama satu bulan terakhir banyaknya penjualan sepeda motor pada lima dealer di Kota A adalah 28, 22, 25, 20, dan 25. Tentukan Median data tersebut.

Penyelesaian:

Diperoleh $j = \frac{(5+1)}{2} = 3$, sehingga Median = $x_j = x_3$. Median terletak di data ke-3 setelah data diurutkan.

20 22 **25** 25 28

Contoh 4-6

Sebagai contoh, seorang peneliti melakukan pencatatan terhadap bayi yang lahir selama setahun terakhir pada delapan Desa sampel. Hasil pencatatannya adalah 33, 28, 31, 41, 36, 25, 30, dan 32. Tentukan Median data tersebut.

Penyelesaian:

Diperoleh $j = \frac{8}{2} = 4$, sehingga $x_j = x_4$ dan $x_{j+1} = x_5$. Median terletak di antara data ke-4 dan ke-5 setelah data diurutkan.

25 28 30 **31** **32** 33 36 41

Median = $\frac{(x_4+x_5)}{2} = \frac{(31+32)}{2} = 31{,}5$

b. **Median Data Berkelompok**

Median untuk data berkelompok adalah nilai yang letaknya tepat ditengah dengan data yang telah dikelompokkan dalam bentuk distribusi frekuensi. Pada data kelompok nilai informasi atau karakteristik dari masing-masing data tidak dapat diidentifikasi lagi, yang dapat diketahui hanyalah karakter dari kelas atau intervalnya. Akibatnya akan terdapat kesulitan dalam menentukan nilai median yang tepat pada suatu interval kelas. Oleh sebab itu, pendugaan dilakukan dengan cara:

1. Menentukan letak median data berkelompok yaitu $\frac{n}{2}$
2. Mendapatkan nilai Median dengan rumus:

$$Median = L + c\left(\frac{\frac{n}{2} - Fsm}{fm}\right)$$

Keterangan:
L : batas bawah kelas median
c : lebar kelas
n : jumlah data frekuensi
Fsm : frekuensi kumulatif sebelum kelas median
fm : frekuensi data pada kelas median

Contoh 4-7
Sebagai contoh, berdasarkan data net profit yang diperoleh 80 UKM pada Tabel 2.2. tentukanlah median data berkelompok tersebut.

Penyelesaian:

Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.4 Perhitungan Median Data Berkelompok

Kelas ke-i	Interval Kelas	f_i	F
1	31 – 40	2	2
2	41 – 50	3	5
3	51 – 60	5	10
4	61 – 70	13	**23**
5	71 – 80	**24**	47
6	81 – 90	21	68
7	91 – 100	12	80

(Sumber: Penulis, 2023. Diolah)

Sehingga diperoleh, Letak Median = $\frac{n}{2} = \frac{80}{2} = 40$

Karena jumlah data 80, maka median terletak pada data ke 40 dan 41. Data 40 dan 41 berada pada interval kelas 71 – 80.

Diketahui: fm = 24 Fsm = 23 L = 70,5 c =10
n = 80

$$Median = L + c\left(\frac{\frac{n}{2} - Fsm}{fm}\right) = 70{,}5 + 10\left(\frac{40 - 23}{24}\right) = 77{,}58$$

3.2.3. Modus

Modus merupakan nilai yang terjadi paling sering, atau memiliki frekuensi paling tinggi. Suatu kumpulan data dapat mempunyai lebih dari satu modus, namun juga modus tidak selalu ada. Cara menentukan modus yaitu dengan memperhatikan data yang mempunyai frekuensi paling tinggi.

Contoh 4-8
Sebagai contoh, selama satu bulan terakhir banyaknya penjualan sepeda motor pada lima dealer di Kota A adalah 28, 22, 25, 20, dan 25. Tentukan Modus data tersebut.

Penyelesaian:
28 22 **25** 20 **25** ; sehingga Modus = 25

Contoh 4-9
Data dua belas mahasiswa yang diambil secara acak dicatat berapa kali mereka melakukan transaksi melalui e-banking selama sebulan lalu. Datanya yaitu: 2, 0, 3, 1, 2, 4, 2, 5, 4, 0, 1, 4 maka modusnya adalah?

Penyelesaian:
2 0 3 1 **2** **4** **2** 5 **4** 0 1 **4** ; sehingga Modus = 2 dan 4

Modus Data Berkelompok

Dalam menentukan nilai modus data berkelompok langkah awal yang dilakukan adalah menentukan kelas modus yaitu kelas yang mempunyai frekuensi paling tinggi. Setelah kelas modus ditentuskan maka nilai modus dapat diperoleh dengan menggunakan rumus:

$$Modus = L + c\left(\frac{d_1}{d_1 + d_2}\right)$$

Keterangan:

L : batas bawah kelas modus
c : lebar kelas
d_1 : selisih antara frekuensi kelas modus dengan frekuensi kelas sebelumnya
d_2 : selisih antara frekuensi kelas modus dengan frekuensi kelas sesudahnya

Contoh 4-10

Sebagai contoh, berdasarkan data net profit yang diperoleh 80 UKM pada Tabel 2.2. tentukanlah modus data berkelompok tersebut.

Penyelesaian:

Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.5 Perhitungan Modus Data Berkelompok

Kelas ke-i	Interval Kelas	f_i
1	31 – 40	2
2	41 – 50	3
3	51 – 60	5
4	61 – 70	**13**
5	71 – 80	**24**
6	81 – 90	**21**
7	91 – 100	12

(Sumber: Penulis, 2023. Diolah)

Kelas dengan frekuensi tertinggi yaitu kelas ke-5 sebanyak 24 data. Dengan demikian maka diketahui:

L = 70,5 c =10 d_1= 24-13=11 d_2 = 24-21=3

$$Modus = L + c\left(\frac{d_1}{d_1 + d_2}\right) = 70{,}5 + 10\left(\frac{11}{11 + 3}\right) = 78{,}35$$

3.2.4. Ukuran Letak

Ukuran letak menunjukkan pada bagian mana data terletak pada suatu kumpulan data yang telah diurutkan. Ukuran letak meliputi kuartil, desil, dan persentil.

a. Kuartil

Kuartil merupakan ukuran letak yang membagi data menjadi 4 bagian sama besar setelah data diurutkan, setiap bagian dari kuartil sebesar 25%. Rumus mencari letak Kuartil data tidak berkelompok yaitu:

$$Letak\ K_i = \frac{i(n+1)}{4} \quad ; \quad i = 1,2,3$$

Contoh 4-11

Sebagai contoh, seorang peneliti melakukan pencatatan terhadap pengeluaran (dalam juta rupiah) sebelas mahasiswa selama sebulan lalu. Datanya yaitu: 1,5 2,0 1,8 1,6 2,7 0,8 2,5 0,9 2,3 1,3 1,4. Tentukanlah nilai K_1, K_2 dan K_3 data tersebut.

Penyelesaian:

$Letak\ K_1 = \frac{1(11+1)}{4} = 3$

$Letak\ K_2 = \frac{2(11+1)}{4} = 6$

$Letak\ K_3 = \frac{3(11+1)}{4} = 9$

Sehingga, diperoleh letak dan nilai K_1, K_2 dan K_3 adalah:

0,8	0,9	**1,3**	1,4	1,5	**1,6**	1,8	2,0	**2,3**	2,5	2,7
		K_1			K_2			K_3		

Jika diperhatikan, $K_1 = 1,3$ sejajar dengan 25% data, $K_2 = 1,6$ dengan 50%, dan $K_3 = 2,3$ dengan 75% data. Kuartil membagi data menjadi 4 bagian sama besar.

Contoh 4-12

Sebagai contoh, seorang peneliti melakukan pencatatan terhadap bayi yang lahir selama setahun terakhir pada delapan Desa sampel. Hasil pencatatannya adalah 33, 28, 31,

41, 36, 25, 30, dan 32. Tentukan nilai K_1, K_2 dan K_3 data tersebut.

Penyelesaian:

$Letak\ K_1 = \frac{1(8+1)}{4} = 2{,}25$

$Letak\ K_2 = \frac{2(8+1)}{4} = 4{,}50$

$Letak\ K_3 = \frac{3(8+1)}{4} = 6{,}75$

Sehingga, diperoleh letak dan nilai K_1, K_2 dan K_3 adalah:

25 28 30 31 32 33 36 41

K_1 K_2 K_2

$K_1 = X_2 + 0{,}25(X_3 - X_2)$
$= 28 + 0{,}25(30 - 27)$
$= 28{,}75$

$K_2 = X_4 + 0{,}50(X_5 - X_4)$
$= 31 + 0{,}50(32 - 31)$
$= 31{,}50$

$K_3 = X_6 + 0{,}75(X_7 - X_6)$
$= 33 + 0{,}75(36 - 33)$
$= 35{,}25$

Kuartil Data Berkelompok

Nilai Kuartil data berkelompok dapat diperoleh dengan cara:

1. Menentukan Letak Kuartil data berkelompok yaitu: $Letak\ K_i = \frac{i}{4} \times n \quad ; \quad i = 1{,}2{,}3$
2. Mendapatkan nilai Kuartil dengan rumus:

$$K_i = L + c\left(\frac{\frac{i.n}{4} - Fsk}{fk}\right)$$

Keterangan:

K_i : Kuartil ke-i ; i = 1,2,3
L : batas bawah kelas kuartil
c : lebar kelas
n : jumlah data frekuensi

Fsk : frekuensi kumulatif sebelum kelas kuartil

fk : frekuensi data pada kelas kuartil

Contoh 4-13

Sebagai contoh, berdasarkan data net profit yang diperoleh 80 UKM pada Tabel 2.2. tentukanlah K_3 data berkelompok tersebut.

Penyelesaian:

Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.6 Perhitungan Kuartil Data Berkelompok

Kelas ke-i	Interval Kelas	f_i	F
1	31 – 40	2	2
2	41 – 50	3	5
3	51 – 60	5	10
4	61 – 70	13	23
5	71 – 80	24	**47**
6	81 – 90	**21**	68
7	91 – 100	12	80

(Sumber: Penulis, 2023. Diolah)

Sehingga diperoleh, $Letak\ K_3 = \frac{3}{4} \times 80 = 60$

Kuartil terletak pada data ke 60 maka berada pada interval kelas 81 – 90. Diketahui:

fk = 21 Fsk = 47 L = 80,5 c =10 n = 80

$$K_3 = L + c\left(\frac{\frac{i.n}{4} - Fsk}{fk}\right) = 80{,}5 + 10\left(\frac{60 - 47}{21}\right) = 86{,}69$$

b. Desil

Desil merupakan ukuran letak yang membagi data menjadi 10 bagian sama besar setelah data diurutkan, setiap bagian dari

desil sebesar 10%. Rumus mencari letak Desil data tidak berkelompok yaitu:

$$Letak\ D_i = \frac{i(n+1)}{10} \quad ; \quad i = 1,2,\dots,9$$

Contoh 4-14

Sebagai contoh, seorang peneliti melakukan pencatatan terhadap pengeluaran (dalam juta rupiah) sebelas mahasiswa selama sebulan lalu. Datanya yaitu: 1,5 2,0 1,8 1,6 2,7 0,8 2,5 0,9 2,3 1,3 1,4. Tentukanlah nilai D_2 dan D_6 data tersebut.

Penyelesaian:

$Letak\ D_2 = \frac{2(11+1)}{10} = 2,4$

$Letak\ D_6 = \frac{6(11+1)}{10} = 7,2$

Sehingga, diperoleh letak dan nilai D_2 dan D_6 adalah:

0,8 0,9 1,3 1,4 1,5 1,6 1,8 2,0 2,3 2,5 2,7

D_2 D_6

$D_2 = X_2 + 0,4(X_3 - X_2)$
$= 0,9 + 0,4(1,3 - 0,9)$
$= 1,06$

$D_6 = X_7 + 0,2(X_8 - X_7)$
$= 2,0 + 0,2(2,0 - 1,8)$
$= 2,04$

Desil Data Berkelompok

Nilai Desil data berkelompok dapat diperoleh dengan cara:

1. Menentukan Letak Desil data berkelompok yaitu: $Letak\ D_i = \frac{i}{10} \times n \quad ; \quad i = 1,2,\dots,9$

2. Mendapatkan nilai Desil dengan rumus:

$$D_i = L + c\left(\frac{\frac{i.n}{10} - Fsd}{fd}\right)$$

Keterangan:

D_i : Desil ke-i ; i = 1,2,...,9

L : batas bawah kelas desil
c : lebar kelas
n : jumlah data frekuensi
Fsd : frekuensi kumulatif sebelum kelas desil
fd : frekuensi data pada kelas desil

Contoh 4-15
Sebagai contoh, berdasarkan data net profit yang diperoleh 80 UKM pada Tabel 2.2. tentukanlah D_4 data berkelompok tersebut.

Penyelesaian:

Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.7 Perhitungan Desil Data Berkelompok

Kelas ke-i	Interval Kelas	f_i	F
1	31 – 40	2	2
2	41 – 50	3	5
3	51 – 60	5	10
4	61 – 70	13	**23**
5	71 – 80	**24**	47
6	81 – 90	21	68
7	91 – 100	12	80

(Sumber: Penulis, 2023. Diolah)

Sehingga diperoleh, $Letak\ D_4 = \frac{4}{10} \times 80 = 32$

Desil terletak pada data ke 32 maka berada pada interval kelas 71 – 80.

Diketahui:

fd = 24 Fsd = 23 L = 70,5 c =10 n = 80

$$D_4 = L + c\left(\frac{\frac{i.n}{10} - Fsd}{fd}\right) = 70{,}5 + 10\left(\frac{32-23}{24}\right) = 74{,}25$$

c. Persentil

Persentil merupakan ukuran letak yang membagi data menjadi 100 bagian sama besar setelah data diurutkan, setiap bagian dari persentil sebesar 1%. Rumus mencari letak Persentil data tidak berkelompok yaitu:

$$Letak\ P_i = \frac{i(n+1)}{100} \quad ; \quad i = 1{,}2, \ldots, 99$$

Contoh 4-16

Sebagai contoh, seorang peneliti melakukan pencatatan terhadap pengeluaran (dalam juta rupiah) sebelas mahasiswa selama sebulan lalu. Datanya yaitu: 1,5 2,0 1,8 1,6 2,7 0,8 2,5 0,9 2,3 1,3 1,4. Tentukanlah nilai P_{25} dan P_{66} data tersebut.

Penyelesaian:

$Letak\ P_{27} = \frac{27(11+1)}{100} = 3{,}24$

$Letak\ P_{66} = \frac{66(11+1)}{10} = 7{,}92$

Sehingga, diperoleh letak dan nilai P_{27} dan P_{66} adalah:

0,8 0,9 1,3 1,4 1,5 1,6 1,8 2,0 2,3 2,5 2,7

P_{27} P_{66}

$P_{27} = X_3 + 0{,}24(X_4 - X_3)$

$= 1{,}3 + 0{,}24(1{,}4 - 1{,}3)$

$= 1{,}324$

$P_{66} = X_7 + 0{,}92(X_8 - X_7)$

$= 1{,}8 + 0{,}92(2{,}0 - 1{,}8)$

$= 1{,}984$

Persentil Data Berkelompok

Nilai Persentil data berkelompok dapat diperoleh dengan cara:

1. Menentukan Letak Persentil data berkelompok yaitu: $Letak\ P_i = \frac{i}{100} \times n \quad ; \quad i = 1,2, \dots ,99$
2. Mendapatkan nilai Persentil dengan rumus:

$$P_i = L + c\left(\frac{\frac{i.n}{100} - Fsp}{fp}\right)$$

Keterangan:

P_i : Persentil ke-i ; i = 1,2,...,99
L : batas bawah kelas persentil
c : lebar kelas
n : jumlah data frekuensi
Fsp : frekuensi kumulatif sebelum kelas persentil
fp : frekuensi data pada kelas persentil

Contoh 4-17

Sebagai contoh, berdasarkan data net profit yang diperoleh 80 UKM pada Tabel 2.2. tentukanlah P_{75} data berkelompok tersebut.

Penyelesaian:

Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.8 Perhitungan Persentil Data Berkelompok

Kelas ke-i	Interval Kelas	f_i	F
1	31 – 40	2	2
2	41 – 50	3	5
3	51 – 60	5	10
4	61 – 70	13	23
5	71 – 80	24	**47**
6	81 – 90	**21**	68
7	91 – 100	12	80

(Sumber: Penulis, 2023. Diolah)

Sehingga diperoleh, $Letak\ P_{75} = \frac{75}{100} \times 80 = 60$

Desil terletak pada data ke 60 maka berada pada interval kelas 81 – 90.

Diketahui:

fp = 21 Fsp = 47 L = 80,5 c =10 n = 80

$$P_{75} = L + c\left(\frac{\frac{i.n}{100} - Fsp}{fp}\right) = 80{,}5 + 10\left(\frac{60 - 47}{21}\right) = 86{,}69$$

3.3. Ukuran Penyebaran Data

Ukuran penyebaran data merupakan suatu ukuran baik parameter atau statistik untuk mengetahui seberapa besar penyimpangan data dengan nilai rata-rata hitungnya. Ukuran penyebaran membantu mengetahui sejauh mana suatu nilai menyebar dari nilai tengahnya. Beberapa ukuran penyebaran data yaitu *range*, varians, dan standar deviasi.

3.3.1. *Range*

Range merupakan beda antara pengamatan terbesar dan terkecil dalam suatu kumpulan data. Ukuran ini digunakan untuk menunjukkan seberapa jauh data memancar, tetapi tidak menunjukkan keragaman (Sutikno and Ratnaningsih, 2022).

$$Range = Nilai\ Maks - Nilai\ Min$$

Contoh 4-18
Berikut adalah harga tiket kereta api (dalam ribuan rupiah) pada lima jenis kereta api untuk rute Surabaya-Yogyakarta 108, 112, 127, 118 dan 113. Tentukan *Range*-nya.

Penyelesaian:

Range = 127 – 108 = 19

3.3.2. Varians dan Standar Deviasi

Varians dan Standar Deviasi merupakan ukuran penyebaran yang menunjukkan standar penyimpangan atau deviasi data terhadap nilai rata-ratanya. Berbeda dengan *range*,

varians dan standar deviasi menggunakan keseluruhan data dan menunjukkan keragaman (Sutikno and Ratnaningsih, 2022).

a. Varian dan Standar Deviasi Populasi

Berikut adalah rumus yang digunakan untuk menentukan varians populasi (σ^2).

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

Keterangan:
σ^2 : varians populasi
μ : rata-rata populasi
x_i : pengamatan ke-*i*, dimana *i* = 1,2,...,*N*
N : banyaknya data pengamatan dalam populasi.

Sementara standar deviasi populasi merupakan akar kuadrat dari varians populasi, yang menunjukkan standar penyimpangan data terhadap nilai rata-ratanya. Standar deviasi populasi (σ) dapat dirumuskan dengan:

$$\sigma = \sqrt{\sigma^2} \quad \text{atau} \quad \sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

Contoh 4-19

Sebagai contoh, berikut adalah keuntungan bersih (dalam miliar rupiah) enam perusahaan di kota X pada tahun 2022: 7, 5, 9, 7, 8, dan 6. Hitung varians dan simpangan baku dari populasi nilai-nilai tersebut.

Penyelesaian:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{7 + 5 + \cdots + 6}{6} = \frac{42}{6} = 7$$

Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.9 Perhitungan Varians Data Populasi

i	x_i	$(x_i - \mu)$	$(x_i - \mu)^2$

1	7	0	0
2	5	-2	4
3	9	2	4
4	7	0	0
5	8	1	1
6	6	-1	1
∑	42	0	10

(Sumber: Penulis, 2023. Diolah)

Sehingga diperoleh varians populasi,

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} = \frac{0 + 4 + \cdots + 1}{6} = \frac{10}{6} 1{,}667$$

Dan standar deviasi populasi,

$$\sigma = \sqrt{\sigma^2} = \sqrt{1{,}667} = 1{,}29$$

b. Varian dan Standar Deviasi Sampel

Apabila dari suatu populasi diambil sampel berukuran *n*, maka varians sampelnya dapat dirumuskan dengan:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \quad \text{atau} \quad s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)}$$

Keterangan:
s^2 : varians sampel
$\bar{x}$: rata-rata sampel
x_i : pengamatan ke-*i*, dimana *i* = 1,2,...,*n*
n : banyaknya data pengamatan dalam sampel.

Sementara standar deviasi sampel dapat dirumuskan dengan:

$$s = \sqrt{s^2} \quad \text{atau} \quad s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \quad \text{atau}$$

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

Contoh 4-20

Sebagai contoh, seorang peneliti melakukan pencatatan terhadap pengeluaran (dalam juta rupiah) sebelas mahasiswa selama sebulan lalu. Datanya yaitu: 1,5 2,0 1,8 1,6 2,7 0,8 2,5 0,9 2,3 1,3 1,4. Tentukanlah varians dan standar deviasi data sampel tersebut.

Penyelesaian:

Dalam menentukan varians data sampel dapat menggunakan satu dari dua rumus yang ada. Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.10 Perhitungan Varians Data Sampel

i	x_i	x_i^2
1	1,5	2,25
2	2,0	4,00
3	1,8	3,24
4	1,6	2,56
5	2,7	7,29
6	0,8	0,64
7	2,5	6,25
8	0,9	0,81
9	2,3	5,29
10	1,3	1,69
11	1,4	1,96
$\sum$	18,8	35,98

(Sumber: Penulis, 2023. Diolah)

Diketahui:

$\sum_{i=1}^{n} x_i^2 = 2{,}25 + 4{,}00 + \cdots + 1{,}96 = 35{,}98$

$(\sum_{i=1}^{n} x_i)^2 = (1{,}5 + 2{,}0 + \cdots + 1{,}4)^2 = (18{,}8)^2 = 353{,}44$

$n = 11$

Sehingga diperoleh varians sampel,

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)} = \frac{(11)(35{,}98) - 353{,}44}{11(11-1)} = \frac{42.34}{110}$$
$$= 0{,}385$$

Dan standar deviasi sampel,

$$s = \sqrt{s^2} = \sqrt{0{,}385} = 0{,}620$$

c. Varian dan Standar Deviasi Data Berkelompok

Varian data berkelompok hampir sama dengan varians data tidak berkelompok, namun dikalikan dengan frekuensi setiap kelasnya. Varians data berkelompok dirumuskan dengan:

$$s^2 = \frac{\sum_{i=1}^{k} f_i(x_i - \bar{x})^2}{\left(\sum_{i=1}^{k} f_i\right) - 1}$$

Keterangan:
s^2 : varians data berkelompok
$\bar{x}$: rata-rata sampel
x_i : titik tengah kelas ke-*i*, dimana *i* = 1,2,...,*k*
f_i : frekuensi kelas ke-*i*, dimana *i* = 1,2,...,*k*

k : banyak kelas

Sementara standar deviasi sampel dapat dirumuskan dengan:

$$s = \sqrt{s^2} \quad \text{atau} \quad s = \sqrt{\frac{\sum_{i=1}^{k} f_i(x_i - \bar{x})^2}{\left(\sum_{i=1}^{k} f_i\right) - 1}}$$

Contoh 4-21

Sebagai contoh, berdasarkan data net profit yang diperoleh 80 UKM pada Tabel 2.2. tentukanlah varians dan standar deviasi data berkelompok tersebut.

Penyelesaian:

Agar lebih memudahkan diperlukan tabel pembantu berikut.

Tabel 3.11 Perhitungan Varians Data Berkelompok

i	Interval Kelas	Titik Tengah (x_i)	f_i	$x_i . f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 . f_i$
1	31 – 40	35,5	2	71	1650,39	3300,78
2	41 – 50	45,5	3	136,5	937,89	2813,67
3	51 – 60	55,5	5	277,5	425,39	2126,95
4	61 – 70	65,5	13	851,5	112,89	1467,57
5	71 – 80	75,5	24	1812	0,39	9,37
6	81 – 90	85,5	21	1795,5	87,89	1845,70
7	91 – 100	95,5	12	1146	375,39	4504,68
Jumlah			80	6090	3590,23	16068,72

(Sumber: Penulis, 2023. Diolah)

Diketahui:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} = \frac{71 + 136{,}5 + \cdots + 1146}{2 + 3 + \cdots + 12} = \frac{6090}{80} = 76{,}125$$

Sehingga diperoleh varians data kelompok,

$$s^2 = \frac{\sum_{i=1}^{k} f_i (x_i - \bar{x})^2}{\left(\sum_{i=1}^{k} f_i\right) - 1} = \frac{3300{,}78 + \cdots + 4504{,}68}{80 - 1}$$

$$= \frac{16068{,}727}{79} = 203{,}401$$

Dan standar deviasi data kelompok,

$$s = \sqrt{s^2} = \sqrt{203{,}401} = 14{,}261$$

DAFTAR PUSTAKA

Suharyadi and Purwanto S.K. (2016) *Statistika untuk Ekonomi dan Keuangan Modern Edisi 3 Buku 1*. Jakarta: Salemba Empat.

Sutikno and Ratnaningsih, D.J. (2022) *Metode Statistika I*. Banten: Universitas Terbuka.

Walpole, R.E. (1995) *Pengantar Statistika Edisi ke-3*. Jakarta: Gramedia Pustaka Utama.

BAB 4

UKURAN DISTRIBUSI DAN BILANGAN BAKU

Oleh Putri Anugrah Cahya Dewi, M.Pd.

4.1. Pendahuluan

Dalam analisis statistika, kita berhadapan dengan konsep ukuran distribusi yang mengelompokkan data ke dalam dua bentuk utama yaitu distribusi teoritis bentuk diskrit dan distribusi teoritis bentuk kontinu. Beberapa jenis distribusi teoritis bentuk diskrit yaitu Distribusi Bernoulli, Distribusi Binomial, Distribusi Poisson, Distribusi Hipergeometrik, dan lain sebagainya (Yusi and Idris, 2019). Adapun beberapa jenis distribusi teoritis bentuk kontinu yaitu Distribusi Gamma, Distribusi Eksponensial, Distribusi Chi-Kuadrat, Distribusi Normal, dan lain sebagainya (Yusi and Idris, 2019). Dalam bab ini, akan dibahas dua distribusi teoritis bentuk diskrit yaitu Distribusi Binomial dan Distribusi Poisson, serta dua distribusi teoritis bentuk kontinu yaitu Distribusi Chi-Kuadrat dan Distribusi Normal. Setelah mempelajari bab ini, diharapkan peserta didik dapat memahami mengenai Distribusi Binomial, Distribusi Poisson, Distribusi Normal dan Distribusi Chi-Kuadrat, serta dapat menerapkannya dalam bidang ilmu sosial.

4.2. Distribusi Teoritis Bentuk Diskrit

Distribusi diskrit merujuk pada sebuah kumpulan data atau daftar di mana variabel acak memiliki nilai dengan peluang tertentu (Mc Graw, 2008). Variabel acak dalam distribusi diskrit memiliki nilai-nilai yang berhingga atau dapat dihitung. Contohnya, hasil pelemparan dadu, jumlah anak dalam sebuah keluarga, atau jumlah kejadian tertentu dalam suatu periode waktu. Distribusi Binomial dan Distribusi Poisson merupakan

contoh konkret lainnya, di mana kita dapat menghitung probabilitas masing-masing hasil atau peristiwa yang mungkin terjadi.

4.2.1. Distribusi Binomial

Distribusi Binomial merupakan suatu distribusi probabilitas yang digunakan untuk menganalisis percobaan berulang dengan dua hasil yang mungkin, yaitu sukses atau gagal. Setiap percobaan dianggap saling bebas (independen) dan memiliki probabilitas sukses yang tetap. Distribusi Binomial memberikan kita peluang dari berapa kali sukses terjadi dalam sejumlah percobaan yang telah ditentukan sebelumnya. Sebagai contoh, Distribusi Binomial dapat diterapkan dalam survei opini publik untuk mengukur tingkat persetujuan dan ketidaksetujuan terhadap suatu kebijakan tertentu. Dengan mengumpulkan data sejumlah responden, kita dapat menggunakan distribusi Binomial untuk menghitung persentase penduduk yang mendukung kebijakan tersebut.

Secara umum, probabilitas kejadian sukses p dan kejadian gagal q dengan x kejadian yang diharapkan dari n percobaan dirumuskan sebagai berikut:

$$P(x) = \binom{n}{x} p^x q^{n-x}$$

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

di mana:

x = banyaknya kejadian yang diharapkan, dengan x = 0, 1, 2, 3 ,.., n

p = probabilitas kejadian sukses, dengan $0 \leq p \leq 1$

q = probabilitas kejadian gagal, dengan $q = 1 - p$ dan $0 \leq q \leq 1$

n = banyaknya percobaan, dengan $n \geq 0$

Mean ($\boldsymbol{\mu}$)

Rata-rata hitung Distribusi Binomial dapat dihitung dengan rumus sebagai berikut.

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

Varians $(\boldsymbol{\sigma^2})$

Varians Distribusi Binomial dapat dihitung dengan rumus sebagai berikut.

$$\sigma^2 = np(1-p)$$

Simpangan Baku $(\boldsymbol{\sigma})$

Simpangan baku Distribusi Binomial dapat dihitung dengan rumus sebagai berikut.

$$\sigma = \sqrt{\sigma^2}$$

Contoh 4-1

Sebuah *platform* pembelajaran *online* mengadakan ujian akhir ilmu politik bagi para pesertanya yang terdiri atas 15 pertanyaan. Jika setiap pertanyaan memiliki empat opsi jawaban, maka berapakah probabilitas seorang peserta menjawab benar 10 pertanyaan? Tentukan pula rata-rata seorang peserta menjawab benar, beserta varians dan simpangan bakunya!

Penyelesaian:

n = 15

p = $^1/_4 = 0{,}25$

q = 1 - $^1/_4$ = $^3/_4$ $= 0{,}75$

x = 10

maka probabilitasnya yaitu

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

$$P(x) = \frac{15!}{10!\,(15-10)!} \cdot \left(\frac{1}{4}\right)^{10} \left(\frac{3}{4}\right)^{(15-10)}$$

$$P(x) = \frac{15!}{10!\,5!} \cdot \left(\frac{1}{4}\right)^{10} \left(\frac{3}{4}\right)^{5}$$

$$P(x) = 0{,}00068$$

Jadi, probabilitas seorang peserta akan menjawab benar tepat 10 pertanyaan dari 15 pertanyaan dalam ujian tersebut adalah 0,00068 atau 0,07%.

Sedangkan rata-rata hitung, varians dan simpangan bakunya yaitu

$\mu = np$

= 15.0,25 = 3,75

$\sigma^2 = np(1-p)$

= 15.0,25(1-0,25) = 2,8125

$\sigma = \sqrt{\sigma^2}$

$= \sqrt{2{,}8125} = 1{,}6771$

Jadi, rata-rata jumlah pertanyaan yang dijawab benar adalah 3,75, variansnya adalah 2,8125 dan simpangan bakunya adalah 1,6771.

Contoh 4-2

Sebuah organisasi nirlaba meluncurkan kampanye kesadaran terhadap bahaya merokok di kalangan remaja. Terdapat 20 remaja yang terlibat dalam kampanye tersebut. Sebelum kampanye diluncurkan, estimasi bahwa 30% remaja dapat berhasil berhenti merokok karena kampanye tersebut. Berapa probabilitas dari 20 remaja yang terlibat dalam kampanye, paling banyak 5 orang yang akan berhenti merokok?

Penyelesaian:

n = 20

p = 30% = 0,3

q = 1 - p = 1 – 0,3 = 0,7

$x \leq 5$

maka,

$P(X \leq 5) = P(0) + P(1) + P(2) + P(3) + P(4) + P(5)$

$$P(0) = \frac{20!}{0!\,(20-0)!} \cdot (0{,}3)^0 (0{,}7)^{20} = 0{,}0008$$

$$P(1) = \frac{20!}{1!\,(20-1)!} \cdot (0{,}3)^1 (0{,}7)^{19} = 0{,}0068$$

$$P(2) = \frac{20!}{2!\,(20-2)!} \cdot (0{,}3)^2 (0{,}7)^{18} = 0{,}0278$$

$$P(3) = \frac{20!}{3!\,(20-3)!} \cdot (0{,}3)^3 (0{,}7)^{17} = 0{,}0716$$

$$P(4) = \frac{20!}{4!\,(20-4)!} \cdot (0{,}3)^4 (0{,}7)^{16} = 0{,}1304$$

$$P(5) = \frac{20!}{5!\,(20-5)!} \cdot (0{,}3)^5 (0{,}7)^{15} = 0{,}1789$$

sehingga,

$P(X \leq 5) = P(0) + P(1) + P(2) + P(3) + P(4) + P(5)$

$P(X \leq 5) = 0{,}0008 + 0{,}0068 + 0{,}0278 + 0{,}0716 + 0{,}1304 + 0{,}1789$

$P(X \leq 5) = 0{,}4163$

Jadi, probabilitas dari 20 remaja yang terlibat dalam kampanye, paling banyak 5 orang akan berhenti merokok adalah 0,4163 atau 41,63%.

4.2.2. Distribusi Poisson

Distribusi Poisson merupakan distribusi probabilitas dalam statistika yang digunakan untuk menggambarkan frekuensi kejadian langka dalam interval waktu atau ruang tertentu. Distribusi ini berlaku ketika kita memiliki data diskrit dan tidak terbatas dalam jumlahnya, dengan tingkat kejadian yang relatif rendah, tetapi memiliki peluang terjadi yang konstan. Contohnya meliputi peristiwa seperti bencana alam, kecelakaan lalu lintas, atau tiba-tiba terjadinya telepon yang masuk. Distribusi binomial memiliki keterbatasan pada jumlah, oleh karena distribusi poisson digunakan untuk mengatasinya. Misal terdapat n percobaan dengan sampel besar (n > 50) dan p sangat kecil (p < 0,1), maka dapat digunakan Distribusi Poisson.

Secara umum, probabilitas kejadian sukses p dan kejadian gagal q dengan x kejadian yang diharapkan dari n percobaan dengan jumlah besar dirumuskan sebagai berikut:

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

di mana :

x = banyaknya kejadian yang diharapkan, dengan x = 0, 1, 2, 3 , ….

n = banyaknya percobaan, dengan n ≥ 0

μ = rata-rata hitung

e = bilangan napier atau bilangan Euler, dengan e = 2,71828

Mean $(\boldsymbol{\mu})$

Rata-rata hitung Distribusi Poisson dapat dihitung dengan rumus sebagai berikut.

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

Varians $(\boldsymbol{\sigma^2})$

Varians Distribusi Poisson dapat dihitung dengan rumus sebagai berikut.

$$\sigma^2 = np$$

Simpangan Baku ($\boldsymbol{\sigma}$)

Simpangan baku Distribusi Poisson dapat dihitung dengan rumus sebagai berikut.

$$\sigma = \sqrt{\sigma^2} = \sqrt{\mu}$$

Contoh 4-3

Sebuah kota mengalami rata-rata 2 kecelakaan lalu lintas setiap minggu di persimpangan jalan tertentu. Berapa peluang bahwa dalam minggu tertentu tidak ada kecelakaan yang terjadi di persimpangan jalan tersebut?

Penyelesaian :

μ = 2

x = 0

maka,

$$P(X = 0) = \frac{\mu^x e^{-\mu}}{x!}$$

$$P(X = 0) = \frac{2^0(2{,}71728)^{-2}}{0!} = 0{,}1354$$

Jadi, peluang bahwa dalam minggu tertentu tidak ada kecelakaan yang terjadi di persimpangan jalan tersebut adalah 0,1354 atau 13,54%.

Contoh 4-4

Seorang peneliti memeriksa rata-rata 3 email spam yang diterima oleh seorang pengguna internet dalam sehari. Berapa peluang

bahwa pengguna internet akan menerima paling sedikit 6 email spam dalam satu hari?

Penyelesaian :

$\mu = 3$

$x \geq 6$

$P(X \geq 6) = 1 - P(X < 6)$

$P(X < 6) = P(0) + P(1) + P(2) + P(3) + P(4) + P(5)$

$$P(X = 0) = \frac{3^0(2{,}71728)^{-3}}{0!} = 0{,}0498$$

$$P(X = 1) = \frac{3^1(2{,}71728)^{-3}}{1!} = 0{,}1495$$

$$P(X = 2) = \frac{3^2(2{,}71728)^{-3}}{2!} = 0{,}2243$$

$$P(X = 3) = \frac{3^3(2{,}71728)^{-3}}{3!} = 0{,}2243$$

$$P(X = 4) = \frac{3^4(2{,}71728)^{-3}}{4!} = 0{,}1682$$

$$P(X = 5) = \frac{3^5(2{,}71728)^{-3}}{5!} = 0{,}1009$$

$P(X < 6) = P(0) + P(1) + P(2) + P(3) + P(4) + P(5)$

$P(X < 6) = 0{,}0498 + 0{,}1495 + 0{,}2243 + 0{,}2243 + 0{,}1682 + 0{,}100 = 0{,}9161$

sehingga,

$P(X \geq 6) = 1 - 0{,}9161 = 0{,}0839$

Jadi, peluang bahwa pengguna akan menerima paling sedikit 6 email spam dalam satu hari adalah 0,0839 atau 8,39%.

4.3. Distribusi Teoritis Bentuk Kontinu

Distribusi kontinu merupakan ukuran distribusi yang berkaitan dengan variabel yang mengambil sejumlah nilai tak terbatas di dalam interval tertentu. Distribusi kontinu berkaitan dengan variabel acak kontinu, di mana variabel tersebut dapat mengambil nilai dalam rentang kontinu atau berkelanjutan. Variabel acak dalam distribusi kontinu memiliki nilia-nilai yang tidak dapat dihitung atau tak terhingga melainkan bisa diukur. Contoh variabel acak kontinu meliputi pendapatan individu, kecepatan mobil di jalan raya, tinggi badan, atau waktu tempuh. Jenis distribusi kontinu yang akan dibahas pada bab ini adalah Distribusi Normal dan Distribusi Chi-Kuadrat.

4.3.1. Distribusi Normal (Distribusi Gaussian)

Distribusi Normal, juga dikenal sebagai Distribusi Gauss atau Kurva Bell, merupakan tipe distribusi probabilitas kontinu yang memiliki ciri khas berbentuk kurva simetris seperti lonceng, dengan puncak kurva berada di nilai rata-rata. Dalam distribusi ini, sebagian besar data cenderung berada di sekitar nilai rata-rata, sementara data yang semakin menjauh dari rata-rata semakin jarang muncul. Distribusi Normal memiliki dua parameter penting: rata-rata (μ) dan simpangan baku (σ). Bentuk distribusi ini sepenuhnya ditentukan oleh nilai-nilai parameter ini. Distribusi Normal sangat berguna dalam banyak analisis statistik, seperti uji hipotesis, prediksi, dan pemodelan data.

Dalam ilmu sosial, Distribusi Normal sering digunakan untuk menggambarkan karakteristik seperti tingkat pendidikan, tingkat pendapatan, skor tes, dan banyak variabel lainnya. Dengan memahami Distribusi Normal, para peneliti dapat menginterpretasikan dan menganalisis data dengan lebih baik untuk mengambil kesimpulan yang lebih akurat.

Adapun karakteristik dari Distribusi Normal (Wirawan, 2002) yaitu :

1. Grafik selalu di atas sumbu X dan ujung grafik hanya mendekati sumbu X tanpa bersinggungan

2. Grafik memiliki bentuk simetris terhadap rataan
3. Rata-rata, median, dan modus distribusi normal memiliki nilai yang sama
4. Dalam distribusi normal, sekitar 68% dari data berada dalam satu deviasi standar dari rata-rata, sekitar 95% berada dalam dua deviasi standar, dan sekitar 99.7% berada dalam tiga deviasi standar dari rata-rata.
5. Luas daerah kurva akan sama dengan luas satu persegi empat

Distribusi Normal memiliki fungsi kepekatan normal (*normal destiny function*) sebagai berikut :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(0{,}5\sigma_x^2)(x-\mu_x)^2}, dengan - \infty < X < \infty$$

di mana :

σ = simpangan baku

π = bilangan konstan (π = 3,14159)

e = bilangan Euler (e = 2,71828)

μ = rata-rata hitung

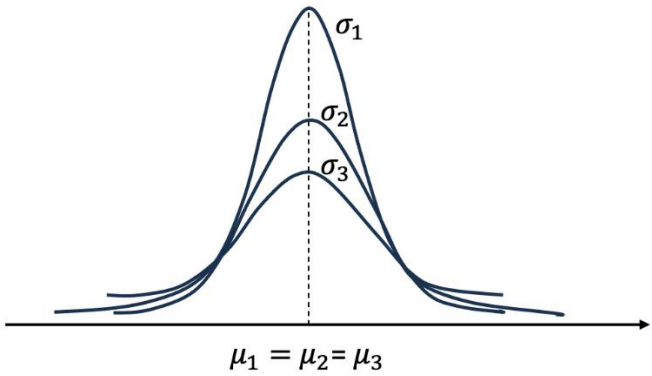
X = variabel acak kontinu

Rata-rata adalah pusat Distribusi Normal. Nilai rata-rata (μ) menunjukkan titik di mana puncak kurva berada. Ketika μ berubah, seluruh kurva akan bergeser ke kiri atau kanan. Jika μ meningkat, kurva akan bergeser ke kanan ; jika μ berkurang, kurva akan bergeser ke kiri. Ini tidak mengubah bentuk kurva, tetapi hanya menggeser posisinya di sepanjang sumbu X.

Sedangkan, simpangan baku (σ) mengontrol seberapa tersebar data di sekitar rata-rata. Semakin besar σ, semakin lebar kurva, dan semakin kecil σ, semakin tajam kurva. Ketika σ berubah, bentuk kurva akan berubah. Jika σ besar maka

akan menghasilkan kurva yang lebih merata dan landai ; jika σ kecil maka akan menghasilkan kurva yang lebih tajam di sekitar puncaknya. Ketika σ sama dengan nol, kurva akan menjadi titik tunggal di nilai rata-rata.

Berikut gambaran dari kurva normal.



Gambar 4.1 Tiga Kurva dengan μ_1= μ_2= μ_3, σ_1< σ_2< σ_3

(Sumber: Penulis, 2023. Diolah)

Bilangan Baku (*Z-Score*)

Z-Skor, juga dikenal sebagai skor z atau standar z, adalah ukuran statistik yang mengukur seberapa jauh suatu nilai individu dari rata-rata dalam satuan simpangan baku (Irianto, 2004). Z-Skor digunakan untuk mengstandarisasi dan membandingkan nilai-nilai yang berasal dari Distribusi Normal.

Distribusi Normal dapat diturunkan ke dalam bentuk distribusi normal baku menggunakan Z-Skor, dengan rumus sebagai berikut.

$$Z = \frac{X - \mu}{\sigma}$$

di mana :

Z = nilai Z-Skor

X = variabel acak kontinu yang terpilih (nilai individu)

μ = rata-rata distribusi normal

σ = simpangan baku distribusi normal

Z-Skor menunjukkan berapa banyak deviasi standar suatu nilai x dari rata-rata μ. Nilai Z-Skor positif menunjukkan bahwa nilai x lebih besar dari rata-rata, sedangkan nilai Z-Skor negatif menunjukkan bahwa nilai x lebih kecil dari rata-rata. Semakin besar nilai absolut Z-Skor, semakin jauh nilai x dari rata-rata dalam satuan simpangan baku.

Contoh 4-5

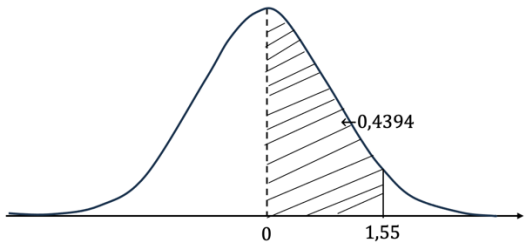
Hitunglah peluang variabel acak normal dengan nilai sebagai berikut.

a. P(0 < Z < 1,55)
b. P(0,2 < Z < 2,3)
c. P(Z < 1,3)

Penyelesaian :

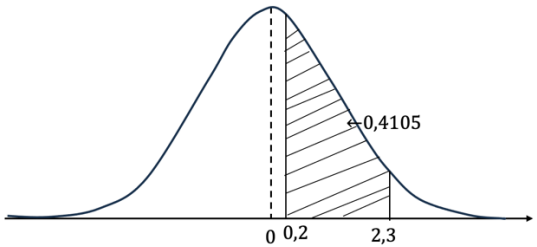
Untuk menyelesaikan soal maka sebaiknya dibuatkan kurvanya terlebih dahulu, kemudian sesuaikan dengan nilai pada tabel normal baku.

a.



Luas daerah dari 0 sampai 1,55 dapat dihitung dengan cara : 1,5 + 0,05, sehingga dari tabel normal baku diperoleh nilai P(0 < Z < 1,55) = 0,4394.

b.

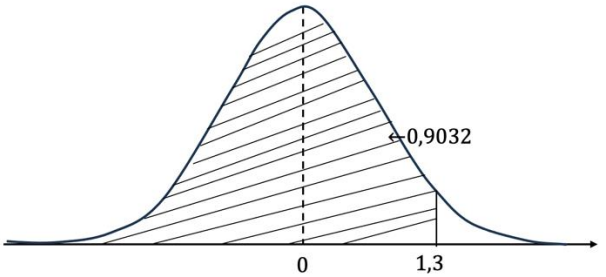


Dari kurva diatas dapat dilihat bahwa nilai P(0,2 < Z < 2,3) = P(0 < Z < 2,3) – P(0 < Z < 0,2).

Luas daerah dari 0 sampai 2,3 dapat dihitung dengan cara : 2,3 + 0,00, sehingga dari tabel normal baku diperoleh nilai 0,4898. Luas daerah 0 sampai 0,2 dapat dihitung dengan cara : 0,2 + 0,00, sehingga dari tabel normal baku diperoleh nilai 0,0793.

Jadi, nilai P(0,2 < Z < 2,3) = 0,4898 – 0,0793 = 0,4105.

c.



Dari kurva diatas dapat dilihat bahwa nilai P(Z < 1,3) = P(Z < 0) + P(0 < Z < 1,3).

Luas daerah sebelah kiri 0 adalah 0,5000. Luas daerah dari 0 sampai 1,3 dapat dihitung dengan cara : 1,3 + 0,00, sehingga dari tabel normal baku diperoleh nilai 0,4032.

Jadi, nilai P(Z < 1,3) = 0,5000 + 0,4032 = 0,9032.

Contoh 4-6

Sebuah ujian yang diambil oleh siswa sekolah menengah memiliki rata-rata skor 75 dan simpangan baku 10. Jika skor ujian memiliki distribusi normal, hitunglah:

a. Peluang bahwa seorang siswa mendapatkan skor di bawah 80.
b. Peluang bahwa seorang siswa mendapat skor antara 70 dan 90.

Penyelesaian:

$\mu = 75$

$\sigma = 10$

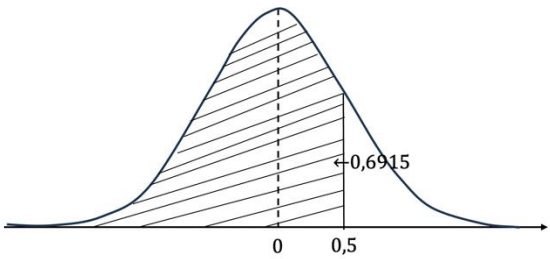
a. P(X < 80)

$$Z = \frac{X - \mu}{\sigma} = \frac{80 - 75}{10} = 0{,}5$$

maka,

P(X < 80) = P(Z < 0,5) = P(Z < 0) + P(0 < Z < 0,5)

P(Z < 0,5) = 0,5000 + 0,1915 = 0,6915



Jadi, peluang bahwa seorang siswa mendapatkan skor di bawah 80 adalah 0,6915 atau 69.15%.

b. P(70 < X < 90)

Untuk x = 70, maka

$$Z = \frac{X - \mu}{\sigma} = \frac{70 - 75}{10} = -0{,}5$$

Untuk x = 90, maka

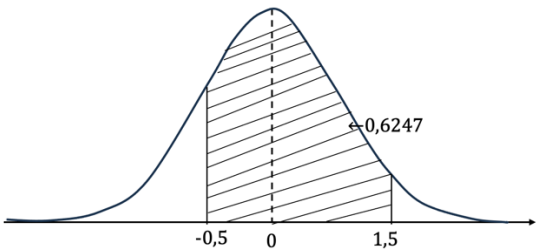
$$Z = \frac{X - \mu}{\sigma} = \frac{90 - 75}{10} = 1{,}5$$

Sehingga diperoleh,

P(70 < X < 90) = P(-0,5 < Z < 1,5)

P(-0,5 < Z < 1,5) = P(0 < Z < 0,5) + P(0 < Z < 1,5)

P(-0,5 < Z < 1,5) = 0,1915 + 0,4332 = 0,6247



Jadi, peluang bahwa seorang siswa mendapat skor antara 70 dan 90 adalah 0,6247 atau 62,47%.

4.3.2. Distribusi Chi-Kuadrat (*Chi-Square*)

Distribusi Chi-Kuadrat (χ^2) merupakan salah satu distribusi probabilitas dalam statistika yang digunakan untuk menggambarkan variasi atau dispersi data. Distribusi Chi-Kuadrat sering digunakan dalam berbagai analisis statistik, terutama dalam pengujian hipotesis terkait varians atau dalam mengukur sejauh mana data menyimpang dari distribusi yang diharapkan (Robert and Brown, 2004). Distribusi Chi-Kuadrat memiliki parameter derajat bebas (*df*), yang menentukan bentuk kurvanya. Semakin besar derajat bebasnya, semakin mendekati distribusi normal. Distribusi Chi-Kuadrat sering digunakan dalam analisis data eksperimental, uji independensi, dan analisis regresi, serta dalam berbagai konteks ilmu sosial dan ilmu alam lainnya. Penting untuk diingat bahwa distribusi Chi-Kuadrat adalah distribusi kontinu dan nilainya selalu positif.

Jika variabel acak saling bebas dan terditribusi normal dengan rata-rata μ dan varians σ^2, maka digunakan rumus berikut :

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma^2}\right)^2, dengan\ df = n$$

Jika variabel acak berukuran n yang berasal dari populasi terdistribusi normal dengan rata-rata μ dan varians σ^2, serta memiliki sampel dengan varians s^2, maka digunakan rumus berikut :

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}, dengan\ df = n - 1$$

Varians ($\boldsymbol{\sigma^2}$)

Secara umum, dengan tingkat kepercayaan $1 - \alpha$, maka diperoleh varians secara keseluruhan adalah antara:

$$\frac{(n-1)^2}{\chi^2\ atas}\ sampai\ \frac{(n-1)^2}{\chi^2\ bawah}$$

Simpangan baku (σ)

Dengan tingkat kepercayaan $1-\alpha$, maka diperoleh simpangan baku adalah antara:

$$\sqrt{\frac{(n-1)s^2}{\chi^2 atas,\, {}^1/_2\,\alpha}}\ sampai\ \sqrt{\frac{(n-1)s^2}{\chi^2 bawah,\, {}^1/_2\,\alpha}}$$

Contoh 4-7

Sebuah lembaga pemerintah mengklaim bahwa rata-rata tingkat kepuasan masyarakat terhadap layanan publik yang disediakan adalah 7 (skala 1-10) dengan simpangan baku 1,5. Sebuah survei dilakukan dengan mengambil sampel dari 5 responden dan mengukur tingkat kepuasan mereka terhadap layanan tersebut: 6, 7, 7.5, 8, dan 8.5. Apakah klaim lembaga pemerintah tentang simpangan baku 1,5 dapat dipercaya?

Penyelesaian :

Karena terdapat sampel, maka digunakan rumus berikut.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}, dengan\ df = n-1$$

Pertama, dicari terlebih dahulu ragam dari sampel.

$$s^2 = \sum \frac{(x-\bar{x})^2}{n-1} = \frac{3{,}7}{4} = 0{,}925$$

Maka,

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$\chi^2 = \frac{4.0{,}925}{1{,}5} = 2{,}467$$

Dengan $\alpha = 5\%$ atau 0,05 dan df = n – 1 = 4 maka dari tabel χ^2 diperoleh

$\chi^2 atas, {}^1/_2\,\alpha\ (\chi^2_{0,025})$ = 0,484

$\chi^2 bawah, {}^1/_2\,\alpha\ (\chi^2_{0,975})$ = 11,14

Karena nilai χ^2 hitung = 2,467 berada di antara interval $\chi^2 atas, {}^1/_2\,\alpha$ dan $\chi^2 bawah, {}^1/_2\,\alpha$, maka klaim lembaga pemerintah tentang simpangan baku 1,5 dapat dipercaya.

DAFTAR PUSTAKA

Irianto, A. (2004) *Statistik Konsep Dasar dan Aplikasinya*. Jakarta Timur: Prenada Media.

Mc Graw, H. (2008) 'Distribusi Probabilitas Normal', p. 35.

Robert, B. and Brown, E. B. (2004) *Probability & Statistics for Engineers & Scientists*.

Wirawan, N. (2002) *Cara Mudah Memahami Statistika 2 (Statistika Inferensia) Untuk Ekonomi dan Bisnis.* Denpasar: Keraras Emas.

Yusi, D. M. S. and Idris, D. U. (2019) *Statistika untuk Ekonomi, Bisnis, & Sosial*. Edited by E. Kurnia. Penerbit Andi.

BAB 5

TEORI PELUANG DAN DISTRIBUSI SAMPLING

Oleh Ketut Queena Fredlina, S.Si., M.Si.

5.1. Pendahuluan

Dalam era informasi dan pengolahan data yang semakin berkembang pesat, pemahaman tentang teori peluang dan distribusi sampling telah menjadi fundamental dalam berbagai bidang ilmu, termasuk statistika, ilmu data, ilmu sosial, dan ilmu alam. Teori peluang merupakan konsep inti yang memungkinkan kita untuk mengukur dan memodelkan ketidakpastian dalam berbagai situasi (Hogg & Tanis, 2019). Sementara distribusi sampling membantu kita dalam membuat inferensi yang dapat diandalkan tentang populasi berdasarkan sampel yang dianalisis (Irianto, 2004). Kedua konsep ini saling melengkapi dan memberikan dasar yang kuat untuk pengambilan keputusan yang berdasarkan bukti dan analisis yang tepat.

Bab ini bertujuan untuk memberikan pemahaman yang kokoh tentang teori peluang dan distribusi sampling serta pentingnya aplikasi keduanya dalam berbagai aspek kehidupan dan penelitian. Pada bab ini akan dibahas konsep dasar dalam teori peluang, termasuk pendekatan perhitungan peluang, hukum peluang, dan peluang bersyarat. Kami juga akan menjelaskan pentingnya distribusi sampling dalam membuat inferensi tentang populasi berdasarkan sampel yang terbatas.

5.2. Pengertian Peluang

Pertama-tama, kita akan menjelajahi pengertian dasar tentang apa itu peluang. Peluang adalah ukuran matematis dari

kemungkinan terjadinya suatu peristiwa. Dalam analisis statistik, peluang membantu kita dalam mengukur seberapa mungkin suatu kejadian akan terjadi berdasarkan informasi yang tersedia (Wackerly et al., 2014).

Peluang biasanya dinyatakan dalam bentuk angka desimal atau pecahan yang berada dalam rentang antara 0 dan 1. Jika peluang suatu kejadian adalah 0, itu berarti kejadian tersebut tidak mungkin terjadi. Jika peluangnya adalah 1, itu berarti kejadian tersebut pasti terjadi. Dalam beberapa kasus, peluang juga dapat dinyatakan dalam bentuk persentase. Jadi, jika peluang suatu kejadian adalah 0.75, ini dapat diartikan sebagai peluang 75% bahwa kejadian tersebut akan terjadi.

Selain itu, dalam beberapa konteks khusus, peluang juga dapat dinyatakan dalam bentuk notasi lain, seperti rasio atau peluang odds. Namun, dalam banyak kasus, representasi paling umum adalah dalam bentuk angka desimal atau pecahan antara 0 dan 1.

Peluang peristiwa A biasanya dilambangkan sebagai $P(A)$, didefinisikan sebagai berikut:

$$P(A) = \frac{\text{jumlah kejadian yang dianggap berhasil}}{\text{total kemungkinan kejadian}}$$

5.3. Pendekatan Perhitungan Peluang

Pendekatan perhitungan peluang adalah suatu metode untuk mengestimasi atau menghitung peluang terjadinya suatu peristiwa berdasarkan pengetahuan yang tersedia mengenai situasi atau eksperimen tertentu. Pendekatan ini melibatkan penggunaan berbagai teknik matematis dan konsep dalam Teori Peluang untuk menghitung peluang dengan menggunakan informasi yang ada (Yusi & Idris, 2019)

Ada beberapa pendekatan yang umum digunakan dalam perhitungan peluang, yaitu pendekatan klasik atau matematis dan pendekatan empiris.

5.3.1. Pendekatan Klasik atau Matematis

Pendekatan ini digunakan ketika setiap hasil dalam ruang sampel memiliki peluang yang sama. Sebagai contoh, dalam pelemparan uang logam yang memiliki dua sisi yaitu angka dan gambar, maka masing-masing sisi memiliki peluang 0,5 atau 1/2. Contoh lain, dalam pelemparan sebuah dadu yang memiliki sisi 1, 2, 3, 4, 5, dan 6, maka masing-masing kemunculan angka akan memiliki peluang 1/6 . Pendekatan ini sangat berguna dalam situasi yang memiliki ruang sampel yang terstruktur dengan baik.

5.3.2. Pendekatan Empiris

Pendekatan empiris melibatkan pengamatan langsung dari hasil eksperimen atau data yang diperoleh dari pengalaman nyata. Metode ini sering digunakan ketika tidak ada informasi matematis yang jelas tentang struktur peluang atau ketika eksperimen yang dilakukan secara berulang untuk mengumpulkan data yang lebih banyak. Pendekatan empiris termasuk pendekatan frekuensi relatif dan pendekatan simulasi. Sebagai contoh, apabila dilakukan pelemparan uang logam sebanyak 100 kali dan misalkan dalam 100 kali percobaan tersebut diperoleh kemunculan sisi angka sebanyak 56 kali, maka peluang yang dihasilkan adalah 56/100.

5.4. Peluang Suatu Kejadian

Dalam konteks Teori Peluang, peluang suatu kejadian merujuk pada ukuran matematis yang menggambarkan seberapa besar kemungkinan suatu peristiwa tertentu terjadi dalam ruang sampel(Irianto, 2004). Suatu kejadian dapat berupa hasil tunggal atau kombinasi dari beberapa hasil dalam ruang sampel.

Apabila suatu percobaan menghasilkan $\boldsymbol{n}$ kejadian berbeda yang mungkin terjadi, dan dari total kejadian tersebut terdapat sejumlah $\boldsymbol{m}$ yang membentuk kejadian $\boldsymbol{A}$ dari ruang sampel $\boldsymbol{S}$, maka **peluang kejadian $\boldsymbol{A}$** adalah:

$$P(A) = \frac{m}{n} = \frac{N(A)}{N(S)}$$

dimana,
$P(A)$ = peluang kejadian A
$m = n(A)$ = banyaknya kejadian A
$n = n(S)$ = banyaknya hasil percobaan yang mungkin terjadi

Dengan kata lain, peluang suatu kejadian adalah perbandingan antara hasil yang kita inginkan (kejadian A) dengan semua kemungkinan hasil dalam percobaan yang dilakukan.

Contoh 5-1:
Pada kejadian pelemparan sebuah dadu, tentukan peluang munculnya angka 5.

Penyelesaian:
Misalkan:
A = kejadian munculnya angka 5 pada pelemparan sebuah dadu
S = ruang sampel = himpunan sisi pada dadu = $\{1, 2, 3, 4, 5, 6\}$

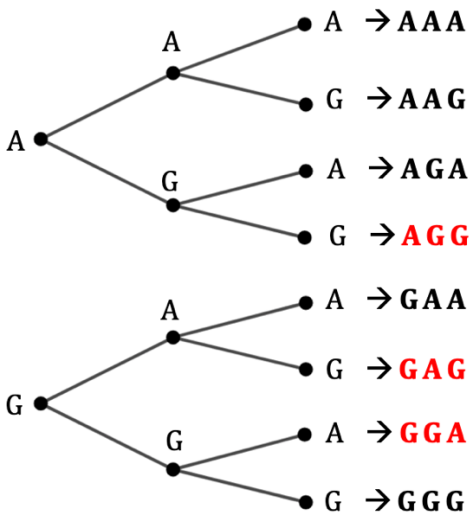
$$P(A) = \frac{n(A)}{n(S)} = \frac{1}{6} \approx 0{,}167$$

Contoh 5-2:
Pada pelemparan 3 buah logam, tentukan peluang munculnya sisi gambar sebanyak 2.

Penyelesaian:
Misalkan:

G = kejadian munculnya sisi gambar sebanyak 2

Cara menentukan ruang sampel dapat dilihat pada Gambar 5.1 berikut :



Gambar 5.1 Penentuan Ruang Sampel Tiga Logam

Peluang terjadinya G:

$$P(G) = \frac{3}{2^3} = \frac{3}{8} = 0{,}375$$

5.4.1. Peluang Komplemen

Peluang komplemen adalah konsep yang erat kaitannya dengan peluang suatu kejadian. Ini merujuk pada peluang bahwa suatu kejadian tertentu tidak terjadi, atau kebalikan dari kejadian yang sedang dipertimbangkan. **Peluang komplemen** atau **peluang tidak terjadinya kejadian *A*** dinotasikan dengan $\boldsymbol{P(\bar{A})}$. Kejadian A dan kejadian bukan A merupakan kejadian yang saling lepas dan interseksinya $(A \cup \bar{A}) = S$ (Yusi & Idris, 2019). Aturan komplemen dinyatakan sebagai berikut :

$$P(A) + P(\bar{A}) = 1$$

sehingga,

$$P(\bar{A}) = 1 - P(A)$$

Contoh 5-3:

Pada pelemparan dua buah dadu, tentukan peluang tidak munculnya mata dadu yang apabila dijumlahkan menghasilkan angka 5.

Penyelesaian:

Misalkan :

A = kejadian munculnya mata dadu dengan jumlah 5
= $\{(1,4), (4,1), (2,3), (3,2)\}$

Cara menentukan ruang sampel dapat dilihat pada Tabel 7.1.

Tabel 5.1 Penentuan Ruang Sampel Dua Dadu

Dadu II / Dadu I	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

(Sumber: Penulis, 2023. Diolah)

Peluang terjadinya A :

$$P(A) = \frac{4}{6^2} = \frac{4}{36}$$

Peluang tidak terjadinya A :

$$P(\bar{A}) = 1 - P(A) = 1 - \frac{4}{36} = \frac{32}{36}$$

5.4.2. Hukum Peluang

Hukum Peluang adalah seperangkat aturan dasar dalam Teori Peluang yang memberikan panduan tentang bagaimana peluang bekerja dan berinteraksi dalam berbagai situasi. Hukum Peluang membantu kita memahami bagaimana menggabungkan, memodifikasi, dan menghitung peluang dari berbagai kejadian yang terkait. Terdapat 2 hukum peluang yang sering digunakan, yaitu **penjumlahan** dan **perkalian** (Wirawan, 2002).

Dalam hukum penjumlahan, apabila A dan B adalah dua kejadian independen, maka peluang salah satu atau keduanya terjadi adalah jumlah dari peluang masing-masing kejadian.

$$P(A \text{ atau } B) = P(A) + (P(B)$$

Sedangkan, apabila A dan B adalah dua kejadian yang tidak independen, maka peluang salah satu atau keduanya terjadi adalah jumlah dari peluang masing-masing kejadian dikurangi peluang terjadi keduanya bersama-sama.

$$P(A \text{ atau } B) = P(A) + (P(B) - P(A \cap B)$$

Contoh 5-4:

a. Dalam pelemparan dadu, tentukan peluang munculnya mata dadu 2 atau 5.
b. Dalam seperangkat kartu remi yang memiliki 52 kartu, tentukan peluang munculnya kartu merah atau kartu gambar (J, Q, dan K).

Penyelesaian:

a. Pada kasus ini, kemunculan angka 2 dan 5 tidak mungkin terjadi secara bersamaan, sehingga kejadian tersebut adalah independen.

$$P(2 \text{ atau } 5) = P(2) + (P(5)$$
$$= \frac{1}{6} + \frac{1}{6}$$

$$= \frac{2}{6} = \frac{1}{3}$$

b. Pada kasus ini, kemunculan kartu berwarna merah dan bergambar dapat terjadi secara bersamaan, sehingga kejadian tersebut tidak independen.

Misalkan : M = Kejadian munculnya kartu merah
G = Kejadian munculnya kartu gambar (J, Q, K)

$$P(M \text{ atau } G) = P(M) + (P(G) - P(M \cap G)$$
$$= \frac{26}{52} + \frac{12}{52} - \frac{6}{52}$$
$$= \frac{32}{52} = \frac{8}{13}$$

Dalam hukum perkalian, apabila A dan B adalah dua kejadian independen, maka peluang keduanya terjadi bersama-sama adalah hasil perkalian dari peluang masing-masing kejadian.

$$P(A \text{ dan } B) = P(A) \times (P(B)$$

Contoh 5-5:

Dalam pelemparan sebuah dadu, tentukan peluang munculnya mata dadu 2 dan 5.

Penyelesaian :

$$P(2 \text{ dan } 5) = P(2) \times (P(5)$$
$$= \frac{1}{6} \times \frac{1}{6}$$
$$= \frac{1}{36}$$

5.4.3. Peluang Bersyarat

Peluang bersyarat merupakan salah satu konsep penting dalam Teori Peluang yang menggambarkan peluang terjadinya suatu kejadian tertentu, dengan asumsi bahwa terdapat kejadian lain yang telah terjadi atau sedang terjadi. Dalam konteks peluang

bersyarat, ingin diketahui peluang suatu kejadian A terjadi, ketika sudah diketahui bahwa kejadian B telah terjadi (Yusi & Idris, 2019). Dalam kondisi yang nyata, tidak semua kejadian-kejadian terjadi secara independen, melainkan saling mempengaruhi. Peluang bersyarat memungkinkan kita untuk mengukur peluang dalam konteks yang lebih realistis dan relevan dengan kondisi sebenarnya.

Peluang bersyarat dari kejadian A terjadi, jika sudah diketahui kejadian B telah terjadi, dinotasikan dengan $P(A|B)$, yang artinya ''Peluang A dengan syarat B terjadi.''

Adapun rumus dari peluang bersyarat :

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

dimana, $P(A \cap B)$ adalah peluang kejadian A dan B terjadi secara bersamaan, dan $P(B)$ adalah peluang kejadian B.

Contoh 5-6:

Suatu perusahaan memiliki 80 karyawan dengan komposisi seperti yang terlihat pada Tabel 7.2 berikut :

Tabel 5.2 Komposisi Karyawan pada Suatu Perusahaan

Bagian	Jenis Kelamin		Jumlah
	Pria (P)	Wanita (W)	
Akuntansi (A)	12	8	20
Logistik (L)	17	9	26
Pemasaran (Pm)	19	15	34
Jumlah	48	32	80

(Sumber: Penulis, 2023. Diolah)

a. Jika seorang karyawan dari bagian Logistik dipilih secara acak, berapa peluang bahwa ia karyawan wanita?
b. Jika seorang karyawan pria dipilih secara acak, berapa peluang bahwa ia berasal dari Bagian Akuntansi?

Penyelesaian:

a. Dalam kasus ini, akan dicari peluang seorang karyawan wanita yang berasal dari bagian Logistik.

$$P(W|L) = \frac{P(W \cap L)}{P(L)}$$
$$= \frac{9}{80} : \frac{26}{80}$$
$$= \frac{9}{26} \approx 0{,}3462$$

b. Dalam kasus ini, akan dicari peluang seorang karyawan dari Bagian Akuntansi yang memiliki jenis kelamin laki-laki.

$$P(A|P) = \frac{P(A \cap P)}{P(P)}$$
$$= \frac{12}{80} : \frac{48}{80}$$
$$= \frac{12}{48} = 0{,}25$$

5.5. Distribusi Sampling

Distribusi Sampling adalah distribusi peluang dari suatu statistik yang dihitung dari sampel acak yang diambil dari suatu populasi. Ini memberikan pandangan tentang variasi yang mungkin terjadi dalam statistik sampel ketika kita mengambil berbagai sampel dari populasi yang sama. Distribusi Sampling sangat penting digunakan dalam statistika inferensial karena dapat membantu dalam memberikan estimasi dan kesimpulan tentang populasi berdasarkan sampel yang terbatas (Irianto, 2004).

Jika dipilih sampel dengan jumlah yang kecil dari populasi besar, maka tidak ada perbedaan yang berarti antara penggunaan metode pengembalian maupun tanpa pengembalian. Pengambilan sampel dengan pengembalian melibatkan kejadian yang independen, yang tidak dipengaruhi oleh hasil sebelumnya, dan kejadian independen tersebut dapat dianalisis dengan mudah.
Terdapat dua jenis distribusi sampling yaitu distribusi sampling rata-rata dan distribusi sampling proporsi.

5.5.1. Distribusi Sampling Rata-Rata

Distribusi Sampling Rata-rata, mengacu pada distribusi peluang dari rata-rata sampel yang diambil dari populasi tertentu. Ini adalah distribusi dari berbagai rata-rata yang mungkin diperoleh dari berbagai sampel yang diambil dari populasi yang sama. Dalam prakteknya, kita menggunakan Distribusi Sampling Rata-rata untuk membuat perkiraan tentang rata-rata populasi berdasarkan rata-rata sampel yang diambil.
Apabila sejumlah n sampel acak diambil dari suatu populasi N yang memiliki rata-rata μ dan standar deviasi σ, maka distribusi sampling rata-rata akan memiliki rata-rata dan standar deviasi (dengan pengembalian) sebesar:

$$\mu = \frac{\sum(x)}{N}$$

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

Contoh 5-7:

Terdapat sekumpulan nilai yang terdiri dari 4, 6, 7, dan 9. Jika ingin diambil 2 sampel secara acak, maka tentukan rata-rata dan standar deviasinya.

Penyelesaian:

Populasi = {4, 6, 7, 9}

Lakukan perhitungan rata-rata dan standar deviasi pada setiap kombinasi sampel. Hasil perhitungan sampel dapat dilihat pada Tabel 5.3.

Tabel 5.3 Perhitungan Sampel Distribusi Sampling Rata-Rata

Sampel ke-	Kombinasi sampel		Rata-rata	Standar deviasi	Peluang
1	4	4	4	0	1/16
2	4	6	5	1	1/16
3	4	7	5,5	1,5	1/16
4	4	9	6,5	2,5	1/16
5	6	4	5	1	1/16
6	6	6	6	0	1/16
7	6	7	6,5	0,5	1/16
8	6	9	7,5	1,5	1/16
9	7	4	5,5	1,5	1/16
10	7	6	6,5	0,5	1/16
11	7	7	7	0	1/16
12	7	9	8	1	1/16
13	9	4	6,5	2,5	1/16
14	9	6	7,5	1,5	1/16
15	9	7	8	1	1/16
16	9	9	9	0	1/16
Rata-rata nilai statistik			**6,5**	**1**	**1**

(Sumber: Penulis, 2023. Diolah)

Berdasarkan perhitungan sampel pada Tabel 7.3, dapat dibuat tabel frekuensi dengan mengumpulkan nilai rata-rata yang

sama. Tabel frekuensi rata-rata sampel dapat dilihat pada Tabel 5.4 berikut:

Tabel 5.4 Frekuensi Rata-Rata Sampel

Rata-rata sampel	Frekuensi (f)	Peluang
4	1	1/16
5	2	2/16
5,5	2	2/16
6	1	1/16
6,5	4	4/16
7	1	1/16
7,5	2	2/16
8	2	2/16
9	1	1/16
Jumlah	**16**	**1**

(Sumber: Penulis, 2023. Diolah)

Jika kita menghitung rata-rata dan standar deviasi populasi, diperoleh:

$$\text{Rata-rata} = \frac{4+6+7+9}{4} = 6{,}5$$

$$\text{Standar deviasi} = \sqrt{\frac{(4-6{,}5)^2+(6-6{,}5)^2+(7-6{,}5)^2+(9-6{,}5)^2}{4}} = 1{,}803$$

Dengan demikian rata-rata dari kombinasi sampel tersebut terkumpul pada suatu nilai yang merupakan rata-rata dari populasi, yaitu 6,5. Meskipun standar deviasi tidak sesuai dengan parameter populasi, perbedaan yang dihasilkan akan relatif kecil apabila digunakan jumlah data yang besar.

5.5.2. Distribusi Sampling Proporsi

Distribusi Sampling Proporsi berkaitan dengan distribusi peluang dari proporsi (persentase) dari suatu kejadian tertentu dalam sampel yang diambil dari populasi. Ini digunakan ketika kita ingin mengambil kesimpulan tentang proporsi populasi berdasarkan proporsi sampel yang diamati. Misalnya, jika kita ingin mengetahui proporsi penduduk yang mendukung kebijakan tertentu, kita dapat mengambil sampel acak dari populasi dan menghitung proporsi mereka yang mendukung kebijakan tersebut. Distribusi Sampling Proporsi membantu kita memahami variasi dalam proporsi sampel dan membantu dalam membuat perkiraan tentang proporsi populasi.

Sama seperti Distribusi Sampling Rata-rata, Distribusi Sampling Proporsi juga memiliki hubungan dengan Distribusi Binomial, yang muncul ketika sampel diambil dari populasi dengan dua hasil yang mungkin, seperti sukses dan gagal.

Contoh 5-8:

Terdapat sekumpulan angka yaitu 1, 2, 3, 5, dan 6. Jika ingin diambil 2 bilangan, tentukan distribusi sampling proporsi untuk bilangan genap. Tentukan juga nilai rata-ratanya kemudian bandingkan dengan nilai rata-rata populasi.

Penyelesaian:

Lakukan perhitungan untuk setiap kombinasi sampel. Berikut merupakan contoh perhitungan untuk salah satu sampel, yaitu (1,2):

$$\text{Proporsi bilangan genap dari } (1,2) = \frac{jumlah\ bilangan\ genap}{total\ bilangan}$$

$$= \frac{1}{2} = 0{,}5$$

Setelah melakukan perhitungan untuk seluruh kombinasi sampel, maka diperoleh hasil perhitungan seperti yang terlihat pada Tabel 5.5.

Tabel 5.5 Perhitungan Sampel Distribusi Sampling Proporsi

Sampel ke-	Kombinasi sampel		Proporsi bilangan genap	Peluang
1	1	1	0	1/25
2	1	2	0,5	1/25
3	1	3	0	1/25
4	1	5	0	1/25
5	1	6	0,5	1/25
6	2	1	0,5	1/25
7	2	2	1	1/25
8	2	3	0,5	1/25
9	2	5	0,5	1/25
10	2	6	1	1/25
11	3	1	0	1/25
12	3	2	0,5	1/25
13	3	3	0	1/25
14	3	5	0	1/25
15	3	6	0,5	1/25
16	5	1	0	1/25
17	5	2	0,5	1/25
18	5	3	0	1/25
19	5	5	0	1/25
20	5	6	0,5	1/25
21	6	1	0,5	1/25
22	6	2	1	1/25
23	6	3	0,5	1/25
24	6	5	0,5	1/25
25	6	6	1	1/25
Rata-rata nilai statistik			**0,4**	**1**

(Sumber: Penulis, 2023. Diolah)

Perhitungan untuk populasi :

$$\text{Proporsi bilangan genap} = \frac{jumlah\ bilangan\ genap\ pada\ populasi}{N}$$
$$= \frac{2}{5} = 0{,}4$$

Dari sini terlihat bahwa nilai distribusi sampling proporsi memiliki nilai yang sama dengan parameter populasi yaitu 0,4.

DAFTAR PUSTAKA

Hogg, R. V., & Tanis, E. A. (2019). *Probability and Statistical Inference*. Pearson.

Irianto, A. (2004). *Statistik Konsep Dasar dan Aplikasinya*. Prenada Media.

Wackerly, D., Mendenhall, W., & Scheaffer, R. L. (2014). *Mathematical Statistics with Applications*. Cengage Learning.

Wirawan, N. (2002). *Cara Mudah Memahami Statistika 2 (Statistika Inferensia) Untuk Ekonomi dan Bisnis*. Keraras Emas.

Yusi, D. M. S., & Idris, D. U. (2019). *Statistika untuk Ekonomi, Bisnis, & Sosial.* (E. Kurnia, Ed.). Penerbit Andi.

BAB 6

STATISTIKA INFERENSIAL DAN PENGUJIAN HIPOTESIS

Oleh: Dr. Darham, S.Pd., M.P.d

6.1. Pendahuluan

Statistika dapat dibedakan berdasarkan tujuan pengolahan data dibedakan menjasi dua jenis yaitu statistika deskriptif dan statistika induktif (inferensial), (Ahyar et al., 2020; Hamzah et al., 2016; Heryana et al., 2023; Kurniawan & Puspitaningtyas, 2016; Priadana & Sunarsi, 2021; Sahir, 2022). Statistika deskriptif biasanya hanya menggambarkan atau sebatas mendeskripsikan data yang telah dikumpulkan peneliti menjadi sebuah informasi yang lengkap. Sedangkan statistika induktif (inferensi) adalah metode untuk mengetahui populasi berdasarkan suatu sampel dengan cara membuat induksi atau penarikan kesimpulan dari karakteristik populasi. Penarikan kesimpulan ini dibuat berdasarkan dugaan (estimasi) yang tahap selanjutnya dilakukan pengujian hipotesis.

Secara proses pengerjaannya statistik inferensial adalah sebuah proses keberlanjutan dari statistik deskriptif. Oleh karena itu, sebelum kita lanjut memahami statistik inferensial ada baiknya peneliti mempelajari dan memperdalam konsep statistik deskriptif, (Pasaribu et al., 2021). Pada bab ini akan mencoba menjelaskan satistika inferensial dan pengujian hipotesis mulai dari jenis-jenis satistika inferensial, kegiatan, ruang lingkup dan fungsinya hingga langkah-langkah penentuan pengujian hipotesis.

6.2. Statistika Inferensial

Statistika Inferensial dapat digunakan untuk melihat hubungan antara variabel serta melihat proses generalisasi secara luas. Teknik statistika inferensial biasanya sudah membuat atau sudah menarik suatu kesimpulan dan keputusan berikut pembuktian hipotesisnya.

6.2.1. Jenis-jenis Statistika inferensial

Statistika inferensial menurut (Amruddin et al., 2022; Heryana et al., 2023) dibedakan menjadi dua berdasarkan hubungan antara variabel penelitian, yaitu:

1. Analisis hubungan
   Analisis ini menguji hipotesis asosiatif, dimana variabel dibagi menjadi dua bagian, variabel bebas dan terikat. Contohnya adalah menganalisa kelompok pengrajin batik desa Purwobakti baik sebelum mereka diberikan pelatihan *digital marketing* maupun sesudah diberikan pelatihan (dalam hal ini pengujiannya dilakukan 2 kali tahap atau 2 test).
2. Analisis Komparatif
   Analisis ini dilakukan dengan cara membandingkan dua atau lebih kelompok data, biasanya kelompok yang saling berpasangan dengan kelompok yang tidak berpasangan. Analisis ini berpedoman terhadap perbedaan dua kelompok yang berbeda dalam waktu tertentu.

   Contohnya adalah peneliti melakukan analisa dengan melihat perbedaan antara dua kelompok pengrajin batik wanita pada waktu tertentu baik yang sudah mengikuti pelatihan *digital marketing* maupun kelompok yang belum pernah ikut pelatihan *digital marketing*.

Statistik inferensial juga dapat dilihat menjadi dua perbedaan berdasarkan data yang digunakan dalam penelitiannya, (Ahyar et al., 2020; Heryana et al., 2023; Priadana & Sunarsi, 2021; Sahir, 2022) yaitu:

a. Statistik parametrik,
   Statistik jenis ini menggunakan data penelitian yang berbentuk skala rasio dan juga berbentuk interval. Asumsi yang digunakan dalam distribusi data jenis ini adalah populasinya berbentuk normal. Statistik jenis ini dapat diuji menggunakan uji *kolmogorov-smirnov* dengan ukuran pengujiannya seperti: *anova, t-test*, regresi dan korelasi.

b. Statistik nonparametrik,
   Statistik jenis ini menggunakan data yang berbentuk skala ordinal dan juga nominal, sehingga asumsi distribusi data populasinya normal tidak perlu dilakukan (Heryana et al., 2023; Kurniawan & Puspitaningtyas, 2016) Ukurannya yang dapat digunakan dalam penelitian tipe ini antara lain adalah:

   1) ***Test binomial***, adalah tes statistik untuk data dari populasi dua kelompok kelas yang berbentuk nominal, dan sampelnya kecil yaitu (n < 25).
   2) ***Runtest***, adalah tes statistik untuk hipotesis deskriptif dengan jumlah sampel satu dan data berbentuk ordinal.
   3) **Test run wald-wolfowitz**, merupakan sebuah tes yang berguna untuk menguji signifikansi hipotesis komparatif dari dua sampel independen, datanya berbentuk ordinal, dan disusun dalam bentuk *run*.
   4) ***Chikuadrat***, adalah tes statistik untuk hipotesis deskriptif dengan data dari populasi dua atau lebih kelas, dan berbentuk nominal, serta sampel penelitiannya besar
   5) ***Chikuadrat dua sampel***, adalah tes statistik untuk hipotesis komparatif antar dua sampel yang besar datanya dan berbentuk nominal.
   6) ***Chikuadrat k sampel***, adalah tes statistik untuk menguji hipotesis komparatif yang lebih dari dua sampel data nominal.
   7) ***Test kolmogorov-smirnov dua sampel***, adalah tes statistik yang menguji hipotesis komparatif dari dua sampel independen yang berbentuk ordinal serta

tersusun dalam tabel distribusi frekuensi kumulatif.

8) ***Test median***, adalah tes statistik untuk menguji tingkat signifikansi hipotesis komparatif dari dua sampel independen dengan datanya yang berbentuk nominal/ordinal.
9) **Median extestion**, untuk menguji hipotesis komparatif media k dari sampel independen penelitian jika datanya berbentuk ordinal.
10) **Signtest**, merupakan sebuah tes hipotesis komparatif dari dua sampel yang berkorelasi, datanya berbentuk ordinal.
11) **Koefisien kontingensi**, digunakan untuk menghitung hubungan antar variable, datanya berbentuk nominal.
12) ***Test friedman***, merupakan sebuah tes yang menguji sebuah hipotesis komparatif dari k sampel berpasangan dari data yang berbentuk ordinal. Jika data mentahnya berbentuk interval atau bentuk rasio harus diubah menjadi ordinal.
13) **Test cochran**, adalah tes statistik untuk menguji sebuah hipotesis komparatif dari k sampel berpasangan yang datanya berbentuk nominal dan frekuensi dikotomi.
14) **Mc Nemar Test**, untuk menguji hipotesis komparatif dua sampel berkorelasi, datanya berbentuk nominal.
15) **Wilcoxon match pairs test**, untuk menguji signifikansi hipotesis komparatif dua sampel yang berkorelasi dan datanya berbentuk ordinal.
16) **Fisher exact probability test**, untuk menguji tingkat signifikansi hipotesis komparatif dua sampel kecil independen dan datanya berbentuk nominal.
17) **Korelasi *spearmanrank***, adalah tes uji statistik yang mencari hubungan atau menguji signifikansi hipotesis asosiatif penelitian dari data yang berbentuk ordinal dan sumber data antar variabelnya tidaklah harus sama.
18) **Korelasi kendalltau**, tes untuk mencari hubungan dan menguji hipotesis antara dua variabel atau lebih, datanya berbentuk ordinal.
19) **Analisis varian satu jalan kruskal-walls**, biasanya

untukmenguji hipotesis dari k sampel independen jika datanya berbentuk ordinal. Jika ditemukan data berbentuk interval atau rasio, perlu diubah ke bentuk data ordinal.

20) **Mann-Whitney U-Test**, untuk menguji tingkat signifikansi dari hipotesis komparatif dua sampel independen apabiladatanya berbentuk ordinal.

Dapat dilihat dari penjabaran sebelumnya statistik inferensial semua testtergantung pada asumsiserta jenis data penelitiannya, (Heryana et al., 2023; Kurniawan & Puspitaningtyas, 2016). Adapun perbedaan masing-masing test secara garis besar dapat dilihat pada tabel 6.1.

Tabel 6.1 Perbedaan Statistika Parametrik dan Nonparametrik

	STATISTIKA PARAMETRIK	STATISTIK NONPARAMETRIK
Deskriptif		
Asumsi distribusi data	Normal	-
Asumsi varians	Homogen	-
Jenis data	Rasio/Interval	Ordinal/Nominal
Ukuran sentral data	Mean	Median
Uji		
Korelasi	Pearson dan Regresi	*Spearman*
Dua Kelompok dan Berbeda	*Independent Sample t Test*	*Mann-Whitney*
Dua Kelompok Lebih dan Berbeda	*Independent One Way* Anova	*Kruskal-Walls*
Berulang dan Dua Kondisi	*Paired Sample t Test*	*Wilcoxon*
Berulang dan Dua Kondisi Lebih	*Repeated One Way* Anova	*Friedman*

Sumber: (Heryana et al., 2023; Kurniawan & Puspitaningtyas, 2016)

6.2.2. Kegiatan, Ruang Lingkup dan Fungsi Statistika Inferensial

Kegiatan statistika inferensia meliputi: pengujian hipotesis, estimasi (menaksir) dan mengambil keputusan. Ruang lingkup pembahasan statistika inferensia dapat meliputi: analisis korelasi, pengujian rata-rata, analisis regresi linier sederhana, analisis varians, analisis kovarians, dan lain sebagainya. Fungsi statistik inferensial adalah untuk melakukan generalisasi hasil penelitian berdasarkan data sampel bagi populasi secara keseluruhan. Fungsi selanjutnya adalah dapat memberikan ramalan (*forecasting*) dari hasil proses analisis untuk pengambilan keputusan di masa depan, membuat taksiran untuk suatu kejadian. Artinya, statistik inferensial dapat memberikan ramalan melalui pendugaan suatu populasi yang kemudian diuji dengan menggunakan hipotesis dengan parameter yang telah ditetapkan (Pasaribu et al., 2021). Statistik inferensial (statistik induktif) mempunyai aturan yang harus dipenuhi sebelum data yang diperoleh dapat dianalisis.

6.3. Pengujian Hipotesis

Pengujian hipotesis merupakan suatu prosedur yang dilakukan peneliti dengan tujuan untuk dapat memutuskan menerima atau menolak hipotesis yang diajukan, (Heryana et al., 2023). Dalam pengujian hipotesis, sebuah keputusan yang diambil belum tentu benar atau salah, oleh karena itu bisa menimbulkan sebuah resiko. Mengetahui besar kecilnya resiko bisa dinyatakan dalam bentuk probabilitas. Dalam statistic inferensia, pengujian hipotesis sangat dibutuhkan, dalam pengujian ini pembuatan keputusan ataupun dalam pemecah persoalan dalam terselesaikan dengan akurat, (Pasaribu et al., 2021).

Adapun besarnya jumlah hipotesis dalam sebuah penelitian yang baik ditentukan dari jumlah masalah yang diteliti, tetapi pada dasarnya hipotesis hanya ada 2 (dua), yaitu: Hipotesis nol (H0) dan Hipotesis alternatip (Ha). Adapun apa dan bagaimana rumusan H0 dan apa rumusan Ha, peneliti dapat

menentukan sendiri, yang penting setiap rumusan hipotesis harus ada alasan atau dasar dari teori yang ada, (Santoso, 2019)

Terdapat dua cara menaksir parameter populasi berdasarkan data sampel, yaitu: *a point estimate* dan *inteval estimate* atau disebut juga *confidence interval*.

a. *A point estimate* adalah suatu taksiran parameter populasi berdasarkan satu nilai data sampel.
b. *Interval estimate* adalah suatu taksiran parameter populasi berdasarkan nilai interval data sampel.

A Point Estimation (Pendugaan Titik)

Pendugaan titik adalah suatu pendugaan data populasi yang diambil dari data sampel penelitian dengan hanya menyebut satu nilai angka tertentu sebagai estimasi untuk parameter yang tidak diketahui, (Pasaribu et al., 2021). Pada pendugaan titik ada dua sifat, pertama memiliki nilai harapan penduga yang harus sama dengan parameter yang akan ditaksir. Kedua, memiliki harapan penduga yang harus mempunyai variasi minimum, setiap penduga titik adalah variable random/mempunyai variasi terkecil dari penaksir titik, (Pasaribu et al., 2021).

Interval estimate (Pendugaan Selang atau Interval)

Pendugaan selang atau interval adalah suatu pendugaan terhadap parameter populasi yang mempunyai nilai diantara dua nilai. Pendugaan selang atau interval merupakan suatu penduga yang mempunyai pengukuran yang obyektif tentang derajat kepercayaan pada ketelitian pendugaan. Pendugaan selang akan memberikan nilai-nilai statistik dalam suatu interval dan bukan sebagai nilai tunggal pada pendugaan parameter. Dalam pendugaan ini, untuk melihat besarnya nilai parameter yang diduga namun tidak diketahui menggunakan probabilitas. Pendugaan selang/interval memiliki sifat interval kepercayaan atau interval keyakinan dan dibatasi oleh batas keyakinan atas dan keyakinan bawah. Jadi semakin besar penduga interval, maka semakin kecil selang kepercayaannya, sedangkan semakin kecil penduga interval, maka semakin besar selang kepercayaannya.

Oleh karena itu, semakin panjang selang kepercayaan, maka semakin yakin bahwa selang tersebut memuat parameter yang tidak diketahuinya. saat menduga suatu batas-batas selang, dengan derajat kepercayaan yang diambil dari normal baku Z akan memperoleh nilai z negative dan z positif yang mempunyai sifat harga yang mutlak dan memiliki ukuran yang sama besar dengan luas daerah antara keduanya(1-α).

Menaksir parameter populasi dengan menggunakan *point estimate* akan mempunyai tingkat risiko kesalahan yang lebih tinggi dibandingkan dengan yang menggunakan *interval estimate*. Makin besar interval taksirannya maka akan semakin kecil kesalahannya. Biasanya, kesalahan taksiran dalam penelitian ditetapkan lebih dahulu, yang pada umumnya digunakan adalah 10%, 5%, dan 1%.

6.3.1. Kemungkinan Kesalahan dalam Pengujian Hipotesis

Pengambilan keputusan dalam uji hipotesis dihadapkan pada dua alternatif (kemungkinan) kesalahan, yaitu:

1. Kesalahan Tipe I (*Type I Error*).
   Kesalahan tipe ini jika menolak hipotesis yang pada hakikatnya adalah benar (seharusnya diterima). Probabilitas kesalahan tipe I ini biasanya disebut dengan risiko alpha (alpha risk) yang dilambangkan dengan simbol α. Tingkat kesalahan dinyatakan dengan α dalam bentuk taraf nyata atau taraf signifikan (*level of significant*), α disebut sebagai tingkat keyakinan (*level of confidence*), karena hal tersebut, kesimpulan yang dibuat adalah benar sebesar α.
2. Kesalahan Tipe II (*Type II error*). Kesalahan yang di perbuat apabila menerima hipotesis yang pada hakikatnya adalah salah (seharusnya ditolak). Probabilitas kesalahan tipe II ini biasanya disebut dengan risiko beta (*beta risk*), dilambangkan dengan simbol β. Tingkat kesalahan ini dinyatakan dengan β yang dalam bentuk penggunaannya disebut sebagai fungsi ciri operasi (operating characteristic function). β disebut juga sebagai kuasa pengujian.

Dua pernyataan yang diperlukan dalam pengujian hipotesis, yaitu:

a. Pernyataan Hipotesis Nol (H0)
   1. Pernyataan yang diasumsikan benar, kecuali ada bukti yang kuat untuk membantahnya;
   2. Selalu mengandung pernyataan “sama dengan”, “tidak ada pengaruh”, “tidak ada perbedaan”; dan
   3. Dilambangkan dengan H0

b. Pernyataan Hipotesis Alternatif (H1)
   1. Pernyataan yang dinyatakan benar jika Hipotesis Nol (H0) berhasil ditolak; dan
   2. Dilambangkan dengan H1 atau H

Pengujian merupakan proses pengambilan keputusan untuk menerima atau menolak hipotesis nol (H0) dengan cara membandingkan nilai a tabel distribusinya (nilai kritis) dengan nilai uji statistiknya, sesuai dengan bentuk pengujiannya, yaitu:

1. H0 diterima, apabila nilai uji statistiknya lebih kecil atau lebih besar dari pada nilai positif atau negatif dari α tabel. Dengan kata lain, nilai uji statistik berada di luar nilai kritis; dan
2. H0 ditolak, apabila nilai uji statistiknya lebih besar atau lebih kecil dari pada nilai positif atau negatif dari α tabel. Dengan kata lain, nilai uji statistik berada di dalam nilai kritis.

Pengambilan kesimpulan analisis statistik berdasarkan pada tingkat signifikansi. Tingkat signifikansi adalah kemampuan untuk digeneralisasikan dengan tingkat kesalahan tertentu. Ada hubungan signifikan, berarti hubungan dapat digeneralisasikan untuk populasi. Ada perbedaan signifikan, berarti perbedaan dapat digeneralisasikan.

6.3.2. Langkah-Langkah Penentuan Pengujian Hipotesis

Dalam ilmu statistik terdapat dua macam statistik inferensial yang dapat digunakan untuk menguji hipotesis, yaitu statistik parametrik dan statistik non-parametrik. Keduanya dikerjakan dengan menggunakan data sampel yang berasal dari sumber data random. Statistik parametrik cenderung digunakan pada data interval dan rasio dengan syarat data yang diambil berdistribusi normal. Statistik nonparametrik digunakan untuk data nominal/diskrit dan ordinal namun tidak perlu syarat data harus berdistribusi normal seperti pada statistik parametrik.

Penentuan teknik statistik yang sesuai untuk pengujian hipotesis dalam penelitian dapat dilakukan dengan melihat data yang tersedia apakah dalam bentuk ordinal atau nomial serta apakah hipotesis yang dibangun adalah hipotesis deskriptif, komparatif atau asosiatif. Pada tabel 6.2 berikut akan menunjukkan teknik statistik yang sebaiknya digunakan untuk pengujian hipotesis

Tabel 6.2 Panduan Teknik Statistik Untuk Pengujian Hipotesis.

Bentuk Data	Hipotesis					
	Deskriptif (satu peubah)	komparatif (dua sampel)		komparatif (> dua sampel)		Asosiatif
		Related	Independen	Related	Independen	
Nominal	Binomial	Mc Nemar	Fisher Exact Probability	$\chi 2$ for k sample	$\chi 2$ for k sample	Contingency Coefficient C
	$\chi 2$ One Sample		$\chi 2$ Two Sample	Cochran Q		
Ordinal	Run Test	Sign test	Median test	Friedman Two Way-Anova	Median Extension	Spearman Rank Correlation
		Wilcoxon matched parts	Mann-Whitney U test		Kruskal-Wallis One Way Anova	Kendall Tau
			Kolmogorov Smirnov			
			Wald-Woldfowitz			
Interval Rasio	T test	T-test of* Related	T-test of* independent	One-Way Anova*	One-Way Anova*	Pearson Product Moment *
				Two Way Anova*	Two Way Anova*	Partial Correlation*
						Multiple Correlation*

Sumber: (Pasaribu et al., 2021)

Langkah-langkah umum pada Uji Hipotesis adalah:

1. Tentukan jenis uji/analisis statistik yang akan digunakan (sesuaikan dengan jenis data skala pengukurannya).
2. Rumuskan hipotesis nol dan hipotesis alternatif.
3. Tentukan tingkat signifikansi a, kekuatan uji(1−β), selisih parameter minimum yang hendak dideteksi.
4. Hitung ukuran sampel minimum yang dibutuhkan. Jika perlu lakukan studi pendahuluan untuk memperoleh data yang dibutuhkan namun belum ada untuk perhitungan ukuran sampel. Pilih statistik penguji yang akan digunakan (sesuai dengan uji/analisis).
5. Pilih statistik penguji yang akan digunakan (sesuai dengan uji/analisis)
    a. Tentukan sampel dan lakukan pengumpulan data.
6. Hitung nilai statistik penguji dari data sampel. Periksa apakah nilai ini terletak pada daerah kritis atau tidak.
7. Hitung nilai p dan buat kesimpulan apakah H0 ditolak atau tidak ditolak.
8. Hitung dan sajikan interval konfidensi100 (1−a)% untuk parameter yang dikaji.
9. Jika H0 tidak ditolak dan ukuran sampel yang diperoleh tidak mencukupi ukuran minimum yang dibutuhkan, hitung kembali kekuatan uji yang sesungguhnya.

Penentuan kekuatan uji (1−β) dan selisih parameter minimum yang hendak dideteksi pada langkah ke-3 serta langkah-langkah ke-4, 6, 9, dan 10 tidak termasuk dalam uji hipotesis yang sebenarnya, namun dalam praktik harus dikerjakan harus dikerjakan dalam urutan tersebut sebagai langkahlangkah metode penelitian.

6.3.3. Jenis-jenis Hipotesis

Menurut (Rosalina et al., 2023) ada tiga jenis hipotesis, yaitu:

1. Hipotesis Deskriptif, hipotesis ini merupakan dugaan terhadap nilai satu variabel dalam satu sampel walaupun di dalamnya bisa terdapat beberapa kategori.

Contoh hipotesis deskriptif:

H0 : Adanya kecederungan mahasiswa menggunakan tas bermerek untuk kuliah

Ha : Adanya kecederungan mahasiswa tidak menggunakan tas bermerek untuk kuliah

2. Hipotesis Komparatif, dugaan terhadap perbandingan nilai dua sampel atau lebih. Biasanya hipotesis komparatif ada yang berpasangan atau komparasi berpasangan (related) dalam dua sampel dan lebih dari dua sampel (k sampel) dan juga dalam bentuk komparasi independen dalam dua sampel dan lebih dari dua sampel (k sampel).

   Contoh hipotesis komparatif sampel Berpasangan, komparatif dua sampel

   Ho : Tidak terdapat perbedaan nilai penjualan pisang sale Sumber Rezeki sebelum dan sesudah ada iklan dimedia sosial.

   Ha : Terdapat berbedaan nilai penjualan pisang sale Sumber Rezeki sebelum dan sesudah ada iklan dimedia sosial.

   Contoh hipotesis komparatif Sampel Independen, komparatif tiga sampel

   Ho : Tidak ada perbedaan antara birokrat, akademisi dan pebisnis UMKM dalam memilih partai.

   Ha : Terdapat perbedaan antara birokrat, akademisi dan pebisnis UMKM dalam memilih partai.

3. Hipotesis Asosiatif, adalah dugaan terhadap hubungan antara dua variabel atau lebih

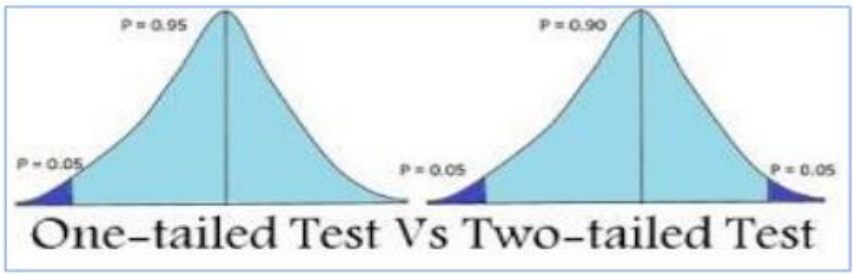
   Contoh :

   Ho : Tidak terdapat hubungan antara jenis profesi dengan jenis cemilan yang disenangi.

   Ha : Terdapat hubungan antara jenis profesi dengan jenis cemilan yang disenangi.

Sedangkan menurut (Santoso, 2019) hipotesis dapat dipisahkan menjadi hipotesis satu ekor dengan dua ekor, aatau satu sisi dengan dua sisi. Hipotesis satu ekor (1-*tailed*) atau dua ekor (2-*tailed*) artinya hipotesis terbagi menjadi dua berdasarkan

arahnya, yaitu hipotesis yang terarah dan tidak terarah (gambar 6.1)



Gambar 6.1One- Tailed Test dan Two- Tailed Test

Sumber: (Santoso, 2019)

Contoh hipotesis satu sisi dan dua sisi:

Diduga uang saku mahasiswa setiap bulannya > Rp. 700.000 (hipotesis satu arah satu ekor, 1-*tailed*, karena sudah menebak salah satu sisi.

Diduga uang saku mahasiswaa setiap bulannya sekitar Rp. 700.000 (hipotesis satu arah dua ekor, 2-*tailed*, karena ada dua kemungkinan, yaitu (lebih besar atau lebih kecil).

6.3.4. Pengujian Hipotesis Deskriptif (Satu sampel)

Hipotesis deskriptif pada data nominal dapat menggunakan uji Test Binomial dan Chi Kuadrat ($\chi 2$) sedangkan jika data berbentuk ordinal maka dapat berbentuk uji yang digunakan adalah Run Test, (Pasaribu et al., 2021) .

Test Binomial

Test Binomial adalah test untuk menguji hipotesis deskriptif (satu sampel) dengan jenis data nominal yang mempunyai dua kelas. Oleh karena itu, uji test binomial sangat sesuai untuk pengujian hipotesis dengan ukuran sampel yang kecil, sedangkan uji chi square sangat tidak mendukung. Pengujian hipotesis menggunakan test binomial dalam satu kelompok data/populasi terdiri dari dua kelas, pertama data berbentuk nominal dan

jumlah sampelnya adalah kecil (<25). Contohnya, kelas jenis kelamin, laki-laki dan perempuan, pendapatan seperti berpenghasilan tinggi dan rendah, prestasi, senior dan junior dan lain sebagainya. Selanjutnya, nilai populasi ini akan diteliti menggunakan sampel yang dari jumlah populasi tersebut.

Chi Kuadrat

Pengujian Chi kuadrat merupakan pengujian hipotesis deskriptif dua kategori kelas atau lebih serta data yang diperoleh berbentuk nominal dengan jumlah sampel yang besar. Hipotesis ini adalah pendugaan atau estimasi terhadap perbedaan rekuensi antara kategori satu dan kategori lain dalam sebuah sampel tentang sesuatu hal.

Adapun rumusnya adalah:

$$X^2 = \sum_{i=1}^{k} \frac{(fo - fh)^2}{fn}$$

Dimana:

x^2 = Chi Kuadrat

fo = Frekuensi Observasi

fh = Frekuensi Diharapkan

Run Test

Uji ini dilakukan untuk mengukur kerandoman populasi berdasarkan data hasil pengamatan melalui data sampel. Kemudian, dari data tersebut dapat diukur banyaknya nilai "run" yang terjadi. Contoh sederhana dalam uji run test adalah pada uang koin. Koin mempunyai dua mata sisi, gambar (G) dan angka (A). Jika sampel kecil, maka penghitungan dapat dijabarkan satu persaatu, sedangkan jika sampel besar, rumusnya adalah:

$$z = \frac{r - \mu_r}{\sigma_r} = \frac{r - \left(\frac{2n_1 n_2}{n_1 + n_2} + 1\right) - 0{,}5}{\sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}}}$$

Kemudian, untuk nilai (mean) atau atau μr dan simpangan baku (σr) (dapat dihitung dengan menggunakan dua rumus berikut:

$$\mu_r = \left(\frac{2n_1n_2}{n_1 + n_2} + 1\right) - 0{,}5$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2}{(n_1 + n_2)^2\,(n_1 + n_2 - 1)}}$$

Pengujian Hipotesis Interval/Rasio

1. Pengujian Rata-Rata, biasanya menggunakan uji Z dengan syarat data berdistribusi normal dengan n yang besar, adapun rumusnya adalah:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \quad z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

2. Pengujian Proporsi, biasanya menggunakan uji Z dengan syarat data berdistribusi normal dengan n yang besar, adapun rumusnya adalah:

$$z = \frac{p - p_0}{\sqrt{p_0q_0/n}}$$

6.3.5. Pengujian Hipotesis Komparatif

Dalam uji hipotesis komparatif pengujian populasi dilakukan dengan membandingkan nilai hitung dari sampel yang telah ditentukan. Sehingga, melalui uji ini kita dapat melihat nilai pengaruh signfikansi peubah independenterhadap peubah dependen.

Uji hipotesis komparatif juga menguji hubungan antara dua sampel atau lebih Pengujian hipotesis komparatif dua sampel berpasangan mempunyai beberapa jenis perhitungan statistik yang dapat diaplikasikan. Masing-masing perhitung bergantung pada jenis data dan kelompok sampelnya seperti yang ditunjukkan dalam tabel 6.2

Mc Nemar Test

Teknik ini digunakan untuk menguji hipotesis komparatif dua sampel yang berkorelasi bila datanya berbentuk niminal atau diskrit. Rancangan penilaian biasanya berbentuk " before after". Jadi hipotesis penelitian merupakan perbandingan antara nilai sebelum dan sesudah yang di dalamnya ada perlakuan. Persamaan dasarnya ditunjukkan sebagai berikut:

$$X^2 = \sum_{i=1}^{K} \frac{(f_o - f_h)^2}{f_h}$$

Dimana:

f o = frekuensi yang diobservasi dalam kategori ke-i
f h= frekuensi yang diharapkan dibawah fo dalam kategori ke-i

Setelah memperoleh nilai Chi Kuadrat tersebut, kita dapat melakukan uji signifikansi dengan menggunakan nilai peubah A dan D, sehingga, rumus sederhana yang dapat digunakan seperti yang ditunjukkan pada rumus berikut:

$$X^2 = \frac{(|A - D| - 1)^2}{A + D}$$

Sign Test (Uji Tanda)

Pengujian hipotesis berikutnya untuk jenis data ordinal adalah sign test atau lebih umunya disebut uji tanda. Penamaan uji tanda ini diberikan mempunyai maksud bahwa semua data yang dianalisis dinyatakan dalam bentuk tanda positif (+) atau negatif (-). Sebagai contoh dalam hal percobaan penerapan jalur satu arah bagi pengemudi mobil dan sepeda motor. Selama percobaan dilakukan hasil tidak dinyatakan dalam bentuk angka atau kuantitatif melainkan dinyatakan apakah percobaan

memberi dampak positif atau negatif terhadap pengurangan kemacetan di jalan.

Fungsi uji tanda dalam percobaan bertujuan melihat pengaruh suatu peubah atau perubahan selama proses eksperimen berlangsung pada keadaan tertentu. Keadaan atau kondisi tersebut menurut John W. Best, adalah: Jika penilaian atas efek peubah atau perlakuan eksperimen tidak dapat diukur, tetapi hanya dapat dinilai dengan sistem juri dalam bentuk performansi baik atau jelek, superior atau inteferior dsb.

6.3.6. Pengujian Hipotesis Asosiatif

Pengujian hipotesis asosiatif menurut (Pasaribu et al., 2021) perlu untuk dilakukan engingat ini akan mempengaruhi setiap peubah apakah masing-masing dari peubah mempunyai hubungan satu sama lain. Terdapat tiga bentuk hubungan antar peubah yaitu peubah berhubungan simetris, sebab-akibat (kausal) dan interaktif/resiprocal (saling mempengaruhi). Caranya adalah dengan menghitung korelasi dari setiap peubah. Arah korelasi dinyatakan dalam bentuk positif atau negatif, sedangkan untuk hubungan antar peubah dinyatakan dalam besarnya koefisien korelasi. Hubungan antara dua peubah atau lebih apabila didapati positif berarti jika salah satu peubah mengalami peningkatan akan meningkatkan nilai peubah lainnya. Sebaliknya, jika nilai suatu peubah menurun maka akan menurunkan nilai peubah lain.

DAFTAR PUSTAKA

Ahyar, H., Maret, U. S., Andriani, H., Sukmana, D. J., Mada, U. G., Hardani, S.Pd., M. S., Nur Hikmatul Auliya, G. C. B., Helmina Andriani, M. S., Fardani, R. A., Ustiawaty, J., Utami, E. F., Sukmana, D. J., & Istiqomah, R. R. (2020). *Buku Metode Penelitian Kualitatif dan Kuantitatif* (Pertama, Issue March). CV. Pustaka Ilmu Group Yogyakarta.

Amruddin, Priyanda, R., Agustina, T. S., Ariantini, N. S., Rusmayani, N. G. A. L., Aslindar, D. A., Ningsih, K. P., & Wulandari, S. (2022). *Bunga Rampai: Metodologi Penelitian Kuantitatif* (F. Sukmawati (ed.)). CV. Pradina Pustaka Grup.

Hamzah, L. M., Awaluddin, I., & Maimunah, E. (2016). *Pengantar Statistika Ekonomi* (Pertama). CV. Anugrah Utama Raharja (AURA).

Heryana, N., Lokollo, L., Ristiyana, R., Effendi, N. I., Manoppo, Y., Khaerani, R., & Seto, A. A. (2023). *Metodologi Penelitian Kuantitatif: dengan Aplikasi IBM SPSS* (N. Mayasari (ed.); Pertama). Get Press Indonesia.

Kurniawan, A. W., & Puspitaningtyas, Z. (2016). *Metode Penelitian Sosial Kuantitatif* (A. W. Kurniawan (ed.); Pertama). Pandiva Buku.

Pasaribu, B., Aji, R. H. S., Utomo, K. W., & Herawati, A. (2021). *Statistika Untuk Ekonomi dan Bisnis* (Cetakan Pe).

Priadana, S., & Sunarsi, D. (2021). *Metode Penelitian Kuantitatif* (Pertama). Pascal Books.

Rosalina, L., Oktarina, R., Rahmiati, & Saputra, I. (2023). *Buku Ajar Statistika* (Eliza (ed.); Pertama). Muharika Rumah Ilmiah.

Sahir, S. H. (2022). *Metodologi Penelitian* (T. Koryati (ed.)). Penerbit KBM Indonesia.

Santoso, I. H. (2019). *Statistik II (Untuk Ilmu Sosial dan Ekonomi)* (R. S. Bahtiar (ed.); Cetakan I).

BAB 7

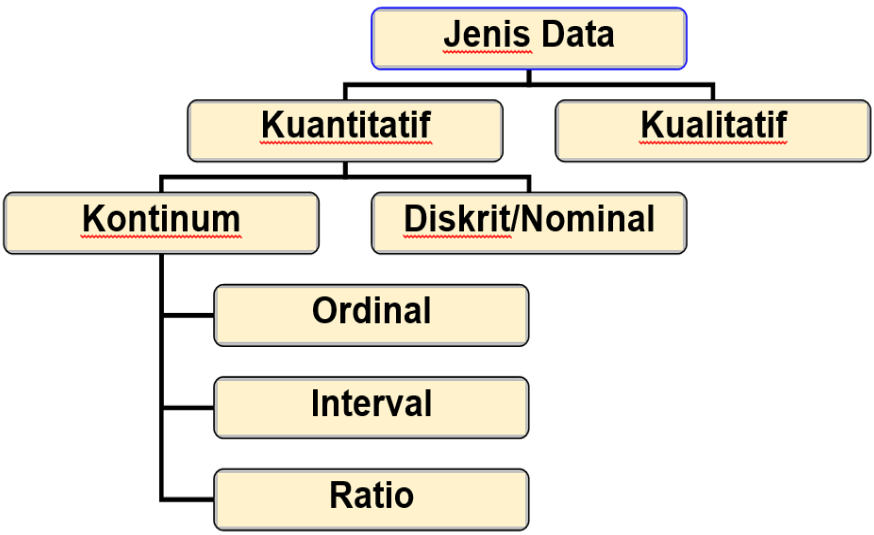
UJI KOMPARATIF SKALA INTERVAL

Oleh Nanang

7.1. Pendahuluan

Data hasil penelitian merupakan bentuk jamak dari datum ialah keterangan atau fakta yang diketahui tentang individu atau kelompok. Sedangkan informasi adalah berita. Contoh data misalnya: *Garut adalah kota Intan*, sementara contoh informasi misalnya: *Di Garut banyak orang yg bernama Intan* (belum tentu benar banyak yang bernama Intan).

Klasifikasi jenis data tampak pada bagan sebagai berikut.



Gambar 7.1 Klasifikasi Jenis Data
(Sumber: Penulis, 2023. Diolah)

1. Data Kualitatif, diperoleh melalui kategorisasi dan dinyatakan dalam bentuk: Kata, Kalimat, atau Gambar.

Contohnya Pengelompokan siswa: Pandai, Sedang, dan Lemah. Siswa SD/MI, SMP/M.Ts., SMA, atau SMK.

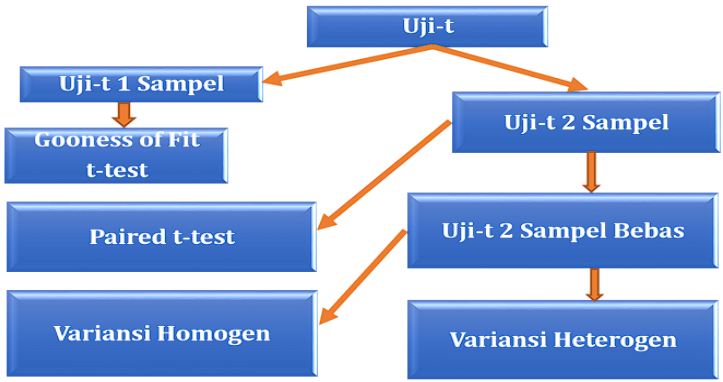
2. Data Kuantitatif terdiri dari:
   a. Data Diskrit digolongkan secara terpisah hasil dari menghitung. Contohnya Jumlah mhs S-2 60 orang. Jumlah RSS: 120 unit dan RS: 100 unit.
   b. Data Kontinum terdiri dari:
      1) Data Ordinal berbentuk ranking. Contohnya Adit ranking 1, Nandia ranking 2, dan Irsan ranking 3.
      2) Data Interval memiliki jarak sama tetapi tdk mempunyai nilai 0 mutlak. Nilai matematika: Adit 80, Nandia 70, dan Irsan 50, Ida 0 (hal ini tidak berarti kemampuan bermatika Ida kosong atau nol).
3. Data Ratio memiliki jarak sama dan nilai nol mutlak. Contohnya Tinggi badan: Adit 140 cm, Nandia 155 cm, Irsan 170 cm. Tinggi kecambah yang belum muncul adalah 0 (hal ini tiggi pohon tersebut benar-benar nol.

Pada bagian ini, khusus akan dibahas untuk data skala interval, yaitu data yang diperoleh dari hasil pengukuran menggunakan alat tes sebagai instrumen alat ukurnya.

7.2. Uji Komparatif Skala Interval

Untuk meneliti suatu fenomena, kita menentukan pengukuran variabel yang diteliti. **D**ari hasil pengukuran skala interval diperoleh data kuantitatif berjarak sama. Uji komparatif merupakan uji membandingkan dua hasil pengukuran variabel terikat. Kelompok yang mendapat perlakuan dibandingkan dengan kolompok mendapat perlakuan lainnya.

Uji komparatif untuk menguji hipotesis masalah praktis sering menggunakan uji-t. Uji-t digunakan ketika informasi mengenai nilai varians populasi tidak diketahui. Uji-t dapat dibagi menjadi 2, yaitu uji-t yang digunakan untuk pengujian hipotesis 1-sampel dan uji-t yang digunakan untuk pengujian hipotesis 2-sampel. Bila dihubungkan dengan kebebasan sampel yang digunakan, maka uji-t dibagi lagi menjadi 2, yaitu uji-t untuk sampel bebas (independent) dan uji-t untuk sampel berpasangan (*paired*). Untuk lebih jelasnya dapat dilihat bagan sebagai berikut.



Gambar 7.2 Penggolongan Uji-t
(Sumber: Penulis, 2023. Diolah)

1. Uji t Satu Sampel (*Goodness of fit t-test*)

Uji perbedaan antara nilai rerata sampel dengan nilai rerata populasi dengan nilai standard tertentu.

Rumus:

$$t_{hitung} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Keterangan:

$\bar{x}$ = nilai rerata sampel

μ_0 = nilai rerata populasi (sebagai standard)

s = simpangan baku sampel

n = ukuran atau besar sampel

Uji dua pihak

$H_0 : \mu_0 = \mu$; $H_1 : \mu_0 \neq \mu$

Ho ditolak untuk:

$t_{hitung} < -\, t_{(1-\alpha/2)\,(df = n-1)}$ atau

$t_{hitung} > t_{(1-\alpha/2)\,(df = n-1)}$

Uji satu pihak

a. $H_0 : \mu_0 \geq \mu$; $H_1 : \mu_0 < \mu$

H0 ditolak untuk: $t_{hitung} < -t_{(1-\alpha)(df = n-1)}$

b. $H_0 : \mu_0 \leq \mu$; $H_1 : \mu_0 > \mu$

H0 ditolak untuk: $t_{hitung} > t_{(1-\alpha)\,(df = n-1)}$

Contoh:

Misalkan kadar nikotin rokok merk **R** diduga lebih tinggi dari 20 mg/batang. Untuk membuktikannya, diambil sampel secara random 10 batang rokok **R**. Selanjutnya diperiksa kadar nikotinnya dengan hasil pemeriksaan: 21, 22, 18, 19, 22, 21, 22, 22, 25, 22. Untuk menguji kebenarannya gunakan α = 0,05 sebagai latihan.

2. Uji-t Dua Sampel

Uji-t dua sampel dibedakan menjadi dua, yaitu:

a. Uji-t dua sampel berhubungan (berpasangan)

Uji-t berpasangan (*paired t-test*) dilakukan pada subjek yang berpasangan ataupun serupa. Misalnya jika kita ingin menguji hasil belajar IPS sebelum menggunakan Lembar Kerja Siswa dengan sesudahnya.

Rumus yang digunakan untuk mencari nilai t dalam ***uji-t berpasangan*** adalah:

$$t_{hitung} = \frac{\bar{d}}{\frac{s}{\sqrt{n}}}$$

Keterangan:

d = Sebelum - Sesudah

= *Paired t test*

$\bar{d}$ = rerata selisih nilai sebelum dan sesudah

s = simpangan baku selisih nilai

$$s = \sqrt{\frac{n\sum d_i^{\,2} - (\sum d_i)^2}{n(n-1)}}$$

n = ukuran atau besar sampel

derajat bebas = n – 1

Hipotesis pada uji-t berpasangan yang digunakan adalah sebagai berikut:

H0: D = 0 (perbedaan antara dua pengamatan sama dengan 0)
Ha: D ≠ 0 (perbedaan antara dua pengamatan tidak sama dengan 0)

Contoh :

Kepada 10 orang lansia diberikan latihan senam. Tekanan darah sebelum dan setelah senam dibandingkan apakah ada perbedaan. Jika hasil pemeriksaan tekanan darah ke-10 orang lansia di bawah ini, apakah ada perbedaan tekanan darah sebelum dan sesudah senam ? (gunakan $\alpha = 0{,}05$)

Subyek	:	1	2	3	4	5	6	7	8	9	10
Sebelum	:	130	128	127	133	134	124	139	132	133	128
sesudah	:	131	129	130	132	129	126	133	128	130	130

(Kerjakan sebagai latihan)

b. Uji-t dua sampel Tidak berhubungan (Independen)

UJI-t 2 sampel dibagi 2 menurut homogenitas variansi kedua sampel, yaitu: (1) variansi kedua sampel homogeny dan (2) variansi kedua sampel tidakhomogen.

1) Bila variansi kedua sampel homogen

$$t_{hitung} = \frac{\overline{X}_1 - \overline{X}_2}{dsg\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ dengan: } dsg = \sqrt{\frac{(n_1 - 1)V_1 + (n_2 - 1)V_2}{n_1 + n_2 - 2}}$$

$\overline{x_1}$ = nilai rerata sampel-1

$v_1 = s_1^2$ = varian sampel-1

$\overline{x_2}$ = nilai rerata sampel-2

$v_2 = s_2^2$ = varian sampel-2

dsg = simpangan baku gabungan kedua sampel

Untuk uji dua pihak: H0 diterima bila :

- $t_{(1-\alpha/2)(df=n1+ n2 - 2)} < t_{hitung} < t_{(1-\alpha/2)\ (df=n1+ n2 - 2)}$

Untuk uji satu pihak: H0 diterima bila :

- $t_{(1-\alpha/2)\ (df=n1+ n2 - 2)} < t_{hitung}$ atau: $t_{hitung} < t_{(1-\alpha/2)\ (df=n1+ n2 - 2)}$

Contoh:

Misal kita ingin mengetahui perbandingan "Hasil belajar siswa dalam pembelajaran Bahasa Indonesia antara yang mendapatkan pembelajaran *Student Team Heroic Leadership* (STHL) dengan siswa yang mendapatkan pembelajaran Ekspositori". Berdasarkan hasil Post-tes terhadap dua kelas (Eksperimen I: STHL dan Eksperimen II: Ekspositori) diperoleh data sebagai berikut.

Tabel 7.1 Data Hasil Posttest Eksperimen I dan Eksperimen II

No	Subjek Eksperimen I	Skor Total	Subjek Eksperimen II	Skor Total
1	EI-18	63	EII-10	56
2	EI-28	63	EII-3	57
3	EI-39	69	EII-33	57
4	EI-15	71	EII-40	57
5	EI-17	72	EII-39	58
6	EI-22	72	EII-34	60
7	EI-7	73	EII-4	64
8	EI-13	73	EII-12	66
9	EI-24	73	EII-38	66
10	EI-29	73	EII-22	70
11	EI-27	74	EII-25	70
12	EI-25	76	EII-26	70
13	EI-26	76	EII-37	70
14	EI-30	78	EII-17	71
15	EI-36	79	EII-20	72
16	EI-38	80	EII-35	73
17	EI-1	81	EII-27	74
18	EI-9	81	EII-6	75

No	Subjek Eksperimen I	Skor Total	Subjek Eksperimen II	Skor Total
19	EI-33	81	EII-21	75
20	EI-10	82	EII-18	77
21	EI-8	85	EII-32	77
22	EI-32	86	EII-31	79
23	EI-14	88	EII-23	80
24	EI-35	89	EII-28	80
25	EI-2	90	EII-1	81
26	EI-4	90	EII-36	82
27	EI-12	90	EII-8	84
28	EI-20	90	EII-9	84
29	EI-31	90	EII-29	85
30	EI-34	90	EII-13	86
31	EI-5	93	EII-24	86
32	EI-23	93	EII-2	87
33	EI-6	95	EII-16	88
34	EI-3	98	EII-19	88
35	EI-11	98	EII-7	89
36	EI-16	98	EII-5	91
37	EI-19	98	EII-11	93
38	EI-21	98	EII-14	95
39	EI-37	98	EII-15	96
30			EII-30	97

(Sumber: Penulis, 2023. Diolah)

a) Uji Normalitas Data Hasil *Posttest* Kelas Eksperimen I

(1) Nilai rata-rata dan simpangan baku

Dari Tabel 9.2 di atas diperoleh nilai rata-rata dan simpangan baku sebagai berikut : $\overline{X}$ = 83,26 dan S = 10,37

(2) Daftar frekuensi observasi dan ekspektasi

(a) Rentang (r) = X_{max} - X_{min} = 98 – 63 = 35

(b) Banyak kelas interval (k) = 1 + 3,3 log 39 = 6,251 (diambil 6)

(c) Panjang kelas (P) ➔ $P = \frac{r}{k} = 35 = 5{,}83$ (diambil 6)

Tabel 7.2 Daftar Frekuensi & Ekspektasi Hasil Posttest ksperimen I

No.	Kelas	Fi	Batas Xi		Nilai Z		Luas Z	Ei	$\frac{(fi-Ei)^2}{Ei}$
1	63 – 68	2	62.5	68.5	-2.00	-1.43	0.0550	2.1450	0.010
2	69 – 74	9	68.5	74.5	-1.42	-0.85	0.1227	4.7853	3.712
3	75 – 80	5	74.5	80.5	-0.84	-0.27	0.1931	7.5309	0.851
4	81 – 86	6	80.5	86.5	-0.27	0.31	0.2281	8.8959	0.943
5	87 – 92	8	86.5	92.5	0.31	0.89	0.1916	7.4724	0.037
6	93 – 98	9	92.5	98.5	0.89	1.47	0.1159	4.5201	4.440
Jumlah		**39**							**9.993**

(Sumber: Penulis, 2023. Diolah)

(3) Mencari nilai Z

$$Z = \frac{Batas\ Kelas - (Rata - rata)}{Simpangan\ Baku}$$

$$Z_1 = \frac{62{,}5 - 83{,}26}{10{,}37} = -2{,}00$$

$$Z_2 = \frac{68{,}5 - 83{,}26}{10{,}37} = -1{,}42$$

$$Z_3 = \frac{74{,}5 - 83{,}26}{10{,}37} = -0{,}84$$

$$Z_4 = \frac{80{,}5 - 83{,}26}{10{,}37} = -0{,}27$$

$$Z_5 = \frac{86{,}5 - 83{,}26}{10{,}37} = 0{,}31$$

$$Z_6 = \frac{92{,}5 - 83{,}26}{10{,}37} = 0{,}89$$

$$Z_7 = \frac{98{,}5 - 83{,}26}{10{,}37} = 1{,}47$$

(4) Mencari luas z (menggunakan daftar z)

L (-2,00 < Z < -1,42) = $|0{,}4772 - 0{,}4222| = 0{,}0550$

L (-1,42 < Z < -0,84) = $|0{,}4222 - 0{,}2995| = 0{,}1227$

L (-0,84 < Z < -0,27) = $|0{,}2995 - 0{,}1064| = 0{,}1931$

L (-0,27 < Z < 0,31) = $|0{,}1064 + 0{,}1217| = 0{,}2281$

L (0,31 < Z < 0,89) = $|0{,}1217 - 0{,}3133| = 0{,}1916$

L (0,89 < Z < 1,47) = $|0{,}3133 - 0{,}4292| = 0{,}1159$

(5) Chi-kuadrat hitung (χ^2 hitung)

$\chi^2_{\text{hitung}} = \sum \frac{(fi - Ei)^2}{Ei} = 0{,}010 + 3{,}712 + 0{,}851 + 0{,}943 + 0{,}037 + 4{,}440 = 9{,}993$

(6) Derajat kebebasan: db = k – 3 = 6 – 3 = 3

(7) Chi-kuadrat tabel (χ^2 tabel) ➔ $\chi^2_{(1-\alpha)}(db) = \chi^2_{(1-0{,}01)}(3) = \chi^2_{(0{,}99)}(3) = 11{,}345$

Karena $\chi^2_{\text{hitung}} = 9{,}993 < \chi^2_{\text{tabel}} = 11{,}345$ maka data hasil *post-tes*t kelas eksperimen I berdistribusi normal.

b) Uji Normalitas Data Hasil *Posttest* Kelas Eksperimen II

(1) Nilai rata-rata dan simpangan baku

Dari Tabel 9.2 di atas diperoleh nilai rata-rata dan simpangan baku sebagai berikut: $\overline{X}$ = 76,65 dan S = 11,76

(2) Daftar frekuensi observasi dan ekspektasi

(a) Rentang (r) = $X_{max} - X_{min}$ = 97 – 56 = 41

(b) Banyak kelas interval (k) = 1 + 3,3 log 41 = 6,28 (diambil 6)

(c) Panjang kelas (P) ➔ $p = \frac{r}{k} = \frac{41}{6} = 6{,}83$ (diambil 7)

Tabel 7.3 Daftar Frekuensi & Ekspektasi Hasil Posttest Eksperimen II

No.	Kelas	Fi	Batas Xi		Nilai Z		Luas Z	Ei	$\frac{(fi-Ei)^2}{Ei}$
1	56 - 62	6	55.5	62.5	-1.80	-1.20	0.0792	3.1680	2.532
2	63 - 69	3	62.5	69.5	-1.20	-0.61	0.1558	6.2320	1.676
3	70 - 76	10	69.5	76.5	-0.61	-0.01	0.2251	9.0040	0.110
4	77 - 83	7	76.5	83.5	-0.01	0.58	0.2230	8.9200	0.413
5	84 - 90	9	83.5	90.5	0.58	1.18	0.1620	6.4800	0.980
6	91 - 97	5	90.5	97.5	1.18	1.77	0.0806	3.2240	0.978
Jumlah		**40**							**7.689**

(Sumber: Penulis, 2023. Diolah)

(3) Mencari nilai Z

$$Z = \frac{Batas\ Kelas - (Rata - rata)}{Simpangan\ Baku}$$

$$Z_1 = \frac{55{,}5 - 76{,}65}{11{,}76} = -1{,}80$$

$$Z_2 = \frac{62{,}5 - 76{,}65}{11{,}76} = -1{,}20$$

$$Z_3 = \frac{69{,}5 - 76{,}65}{11{,}76} = -0{,}61$$

$$Z_4 = \frac{76{,}5 - 76{,}65}{11{,}76} = -0{,}01$$

$$Z_5 = \frac{83{,}5 - 76{,}65}{11{,}76} = 0{,}58$$

$$Z_6 = \frac{90{,}5 - 76{,}65}{11{,}76} = 1{,}18$$

$$Z_7 = \frac{97{,}5 - 76{,}65}{11{,}76} = 1{,}77$$

(4) Mencari luas z (menggunakan daftar z)

L (-1,80 < Z < -1,20) = $|0{,}4641 - 0{,}3849| = 0{,}0792$

L (-1,20 < Z < -0,61) = $|0{,}3849 - 0{,}2291| = 0{,}1558$

L (-0,61 < Z < -0,01) = $|0{,}2291 - 0{,}0040| = 0{,}2251$

L (-0,01 < Z < 0,58) = $|0{,}0040 + 0{,}2190| = 0{,}2230$

L (0,58 < Z < 1,18) = $|0{,}2190 - 0{,}3810| = 0{,}1620$

L (1,18 < Z < 1,77) = $|0{,}3810 - 0{,}4616| = 0{,}0806$

(5) Chi-kuadrat hitung (χ^2 hitung)

$\chi^2_{hitung} = \sum \frac{(fi - Ei)^2}{Ei}$ = 2,532 + 1,676 + 1,110 + 0,413 + 0,980 + 0,978 = 7,689

(6) Derajat kebebasan: db = k – 3 = 6 – 3 = 3

(7) Chi-kuadrat tabel (χ^2 tabel) ➔ $\chi^2_{(1-\alpha)}(db) = \chi^2_{(1-0,01)}(3) = \chi^2_{(0,99)}(3) = 11{,}345$

Karena $\chi^2_{hitung} = 7{,}689 < \chi^2_{tabel} = 11{,}345$ maka hasil postest kelas ekperimen II berdistribusi normal.

Karena kedua kelompok berdistribusi normal, maka langkah selanjutnya adalah menghitung uji homogenitas tes akhir dari kedua kelompok tersebut.

c) Tes Homogenitas Dua Varians Tes Akhir

(1) Mencari nilai F_{hitung}

$$F = \frac{Vb}{Vk} = \frac{(11{,}76)^2}{(10{,}37)^2} = \frac{138{,}2976}{107{,}537} = 1{,}286$$

(2) Menentukan derajat kebebasan

$db_1 = n_1 - 1 = 39 - 1 = 38$ (variansnya $(10{,}37)^2$)

$db_2 = n_2 - 1 = 40 - 1 = 39$ (variansnya $(11{,}76)^2$)

(3) Menentukan nilai F dari daftar

Akan dicari $F_{tabel} = F_{0,01}\,(db_1/db_2) = F_{0,01}\,(38/39)$

Karena di dalam tabel tidak ada maka dicari dengan *interpolasi*

$F_{0,01}\,(30/38) = 2,22$
$F_{0,01}\,(30/40) = 2,20$
} $F_{0,01}\,(30/39) = 2,22 - \frac{1}{2}\,(0,02) = 2,21$……..(1)

$F_{0,01}\,(40/38) = 2,14$
$F_{0,01}(40/40) = 2,11$
} $F_{0,01}\,(40/39) = 2,14 - \frac{1}{2}\,(0,03) = 2,125$……..(2)

Dari persamaan (1) dan (2), maka diperoleh :

$F_{0,01}\,(30/39) = 2,21$
$F_{0,01}(40/39) = 2,125$
} $F_{0,01}\,(38/39) = 2,21 - \frac{1}{10}\,(0,09) = 2,201$

(4) Penentuan homogenitas

Karena $F_{hitung} < F_{tabel}$ atau 1,286 < 2,201 maka varians kedua kelompok homogen.

d) Menguji Perbedaan Rata-Rata Tes Akhir (Uji Satu Pihak)

Karena varians kedua kelompok homogen, maka dilanjutkan ke pengujian perbedaan rata-rata (uji satu pihak).

Langkah-langkahnya adalah :

(1) Merumuskan hipotesis nol dan hipotesis alternatif, sebagai berikut:

Ho = Hasil belajar siswa dalam pembelajaran matematika yang mendapatkan pembelajaran *Student Team Heroic Leadership* (STHL) tidak lebih baik daripada siswa mendapatkan pembelajaran Ekspositori".

Ha = Hasil belajar siswa dalam pembelajaran matematika yang mendapatkan pembelajaran *Student Team Heroic*

Leadership (STHL) lebih baik daripada siswa mendapatkan pembelajaran Ekspositori".

(2) Mencari deviasi standar gabungan

$$S_{gabungan} = \sqrt{\frac{(n_1-1)S_1^{\,2} + (n_2-1)S_2^{\,2}}{n_1+n_2-2}}$$

$$= \sqrt{\frac{(39-1)(10{,}37)^2 + (40-1)(11{,}76)^2}{39+40-2}}$$

$$= \sqrt{\frac{(4086{,}4022) + (5393{,}6064)}{77}}$$

$$= 11{,}10$$

(3) Mencari t_{hitung}

$$t_{hitung} = \frac{\overline{X}_1 - \overline{X}_2}{S_{gab}\sqrt{\frac{n_1+n_2}{n_1 n_2}}} = \frac{83{,}26-76{,}65}{11{,}10\sqrt{\frac{39+40}{(39)(40)}}} = \frac{6{,}61}{11{,}10\sqrt{\frac{79}{1560}}}$$

$$= \frac{6{,}61}{2{,}50} = 2{,}644$$

(4) Menentukan derajat kebebasan

db = $n_1 + n_2 - 2 = 39 + 40 - 2 = 77$

(5) Menentukan nilai t_{tabel} dari daftar

Untuk taraf signifikansi (α) = 1%, maka :

Akan dicari nilai $t_{(1-\alpha)(db)} = t_{(1-0{,}01)(77)} = t_{(0{,}99)(77)}$

$t_{0{,}99(60)} = 2{,}39$

$t_{0{,}99(120)} = 2{,}36$

Jadi,

$$t_{(0{,}99)(77)} = 2{,}39 - \frac{17}{60}(0{,}03) = 2{,}3815$$

(6) Pengujian hipotesis

Kriteria pengujian hipotesis:

Jika : $t_{hitung} < t_{tabel}$ maka Ho diterima dan Ha ditolak.

Dari perhitungan di atas diperoleh t_{hitung} = 2,644 > t_{tabel} = 2,3815 .Ini berarti rata-rata hasil *posttest* kelas eksperimen I

lebih baik daripada kelas eksperimen II. Karena t_{hitung} = 2,644 berada di luar interval $t_{hitung} < t_{tabel}$ maka Ho ditolak dan Ha diterima.

(7) Kesimpulan

Dari pengujian perbedaan rata-rata pada tes akhir dengan menggunakan uji satu pihak, diperoleh kesimpulan : "Hasil belajar siswa dalam pembelajaran matematika yang mendapatkan pembelajaran *Student Team Heroic Leadership* (STHL) lebih baik dibandingkan siswa mendapatkan pembelajaran Ekspositori".

2) Bila variansi kedua sampel tidak homogeny (Uji-t')

Bila kedua sampel berasal dari populasi dengan variansi yang heterogen:

a) Rumus: $t'_{hitung} = \dfrac{\overline{\chi_1} - \overline{\chi_2}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

b) Menentukan kriteria pengujian hipotesis: Ho diterima jika:

$-\dfrac{w_1 t_1 + w_2 t_2}{w_1 + w_2} < t' < \dfrac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$ dengan $w_1 = \dfrac{s_1^{\,2}}{n_1}$; $w_2 = \dfrac{s_2^{\,2}}{n_2}$;

$t_1 = t_\alpha(n_1\text{-}1)$; $t_2 = t_\alpha(n_2\text{-}1)$

Contoh:

Misal kita ingin mengetahui perbandingan "Hasil belajar siswa dalam pembelajaran Bahasa Indonesia antara yang mendapatkan pembelajaran *STAD* dengan siswa yang mendapatkan pembelajaran Konvensional".

Berdasarkan hasil Pre-tes terhadap dua kelas (Eksperimen I: STAD dan Kontrol: Konvensional) diperoleh data sebagai berikut.

Tabel 7.4 Data Hasil Pre-test Kelas Eksperimen dan Kontrol

Subjek Eksperimen	Skor	Subjek Kontrol	Skor
E.1	5	K.1	7
E.2	9	K.2	9
E.3	8	K.3	12
E.4	19	K.4	11
E.5	8	K.5	9
E.6	17	K.6	9
E.7	1	K.7	7
E.8	17	K.8	11
E.9	14	K.9	16
E.10	11	K.10	15
E.11	11	K.11	9
E.12	9	K.12	9
E.13	24	K.13	9
E.14	13	K.14	9
E.15	13	K.15	11
E.16	11	K.16	6
E.17	14	K.17	12
E.18	9	K.18	15
E.19	7	K.19	9
E.20	9	K.20	15
E.21	15	K.21	6
E.22	8	K.22	9
E.23	6	K.23	14
E.24	13	K.24	6
E.25	13	K.25	11
E.26	9	K.26	6
E.27	3	K.27	9
E.28	11	K.28	6
E.29	13	K.29	9
E.30	15	K.30	1
E.31	9	K.31	14
E.32	9	K.32	9
E.33	21	K.33	14
E.34	19	K.34	17
Jumlah (Σ)	**393**		**341**
Rerata ($\overline{X}$)	**11,559**		**10.029**
Simpangan Baku (*s*)	**4,998**		**3.529**

(Sumber: Penulis, 2023. Diolah)

Uji Normalitas

1. Kelas Eksperimen

a. Menentukan nilai rata-rata dan standar deviasi

$\bar{x} = 11{,}559$ dan $s = 4{,}998$

b. Membuat tabel distribusi frekuensi

1). Skor tertinggi = 24

2). Skor terendah = 1

3). Menentukan rentang (r) = data terbesar – data terkecil = 24 – 1 = 23

c. Menentukan banyak kelas interval

k = 1 + 3.3 log n = 1 + log 34 = 6,05diambil k = 6

d. Menentukan panjang kelas interval

$p = \frac{r}{k} = \frac{23}{6} = 3{,}8$ diambil p = 4

Tabel 7.5 Daftar Uji Chi-Kuadrat

Nilai	Fi	Batas xi	Nilai z	Luas z	Ei	$\frac{(f_i - E_i)^2}{E_i}$
		0.5	-2,21			
1 – 4	2			0,0657	2,234	0,025
		4.5	-1,41			
5 – 8	6			0,1917	6,516	0,041
		8.5	-0,61			
9 – 12	11			0,3044	10,350	0,041
		12.5	0,19			
13 – 16	8			0,2636	8,961	0,103
		16.5	0,99			
17 – 19	4			0,1052	3,576	0,050
		19.5	1,59			
20 – 24	2			0,0511	1,738	0,039
		24,5	2,59			

Nilai	Fi	Batas xi	Nilai z	Luas z	Ei	$\frac{(f_i - E_i)^2}{E_i}$
Jumlah	**34**				**Σx^2_{hitung}**	**0,299**

(Sumber: Penulis, 2023. Diolah)

Keterangan :
f_i = frekuensi data
batas x_i = batas kelas
z = transformasi normal standar dari batas kelas dengan rumus

$$z = \frac{x_i - \bar{x}}{s}$$

luas z = luas kelas setiap interval (gunakan tabel z)
E_i = frekuensi ekspektasi dengan rumusE_i = luas z x n

e. Menentukan nilai: $\chi^2_{hitung} = \sum \frac{(f_i - E_i)^2}{E_1} = 0{,}299$

f. Menentukan nilai χ^2_{tabel} dengan α = 5%
db = k – 3 = 6 – 3 = 3
$\chi^2_{tabel} = \chi^2_{(1-0.05)(db)} = \chi^2_{(0.95)(3)} = 7{,}81$

g. Kesimpulan
Karena nilai $\chi^2_{hitung} = 0{,}299 < \chi^2_{tabel} = 7{,}81$ maka data berdistribusi normal.

2. Kelas Kontrol

a. Menentukan nilai rata-rata dan standar deviasi
$\bar{x} = 10{,}029$ dan $s = 3{,}529$

b. Membuat tabel distribusi frekuensi
1). Skor tertinggi = 17
2). Skor terendah = 1
3). Menentukan rentang (r) = data terbesar – data terkecil = 17 – 1 = 16

c. Menentukan banyak kelas interval
k = 1 + 3.3 log n = 1 + log 34 = 6.05d diambil k = 6

d. Menentukan panjang kelas interval

$$p = \frac{r}{k} = \frac{16}{6} = 2{,}6 \text{ diambil p = 3}$$

Tabel 7.6 Daftar Uji Chi-Kuadrat

Nilai	Fi	Batas xi	Nilai z	Luas z	Ei	$\frac{(f_i - E_i)^2}{E_i}$
		0.5	-2,83			
1 – 3	1			0,0239	0,811	0,044
		3,5	-1,94			
4 – 6	5			0,1184	4,025	0,236
		6,5	-1,06			
7 – 9	14			0,2879	9,790	1,811
		9,5	-0,17			
10 – 12	6			0,3286	11,174	2,396
		12,5	0,71			
13 – 15	6			0,1841	6,258	0,011
		15,5	1,6			
16 – 18	2			0,0482	1,640	0,079
		18,5	2,48			
Jumlah	**34**				**Σx^2_{hitung}**	**4,576**

(Sumber: Penulis, 2023. Diolah)

Keterangan :

f_i = frekuensi data

batas x_i = batas kelas

z = transformasi normal standar dari batas kelas

dengan rumus: $z = \frac{x_i - \bar{x}}{s}$

luas z = luas kelas setiap interval (gunakan tabel z)

E_i = frekuensi ekspektasi dengan rumus

E_i = luas z x n

e. Menentukan nilai

$$\chi^2_{hitung} = \sum \frac{(f_i - E_i)^2}{E_1} = 4{,}576$$

f. Menentukan nilai χ^2_{tabel} dengan α = 5%
   db = k – 3 = 6 – 3 = 3
   $\chi^2_{tabel} = \chi^2_{(1-0.05)(db)} = \chi^2_{(0.95)(3)} = 7.81$

g. Kesimpulan
   Karena nilai $\chi^2_{hitung} = 4{,}576 < \chi^2_{tabel} = 7.81$ maka data berdistribusi normal.

   Karena kedua data berdistribusi normal maka dilanjutkan dengan uji homogenitas.

Uji Homogenitas

1. Menentukan nilai F_{hitung} dengan rumus :

$$F = \frac{\text{var } ians_{besar}}{\text{var } ians_{kecil}} = \frac{(s_b)^2}{(s_k)^2}$$

$$F = \frac{(4{,}998)^2}{(3{,}39)^2} = 2{,}174$$

2. Menentukan nilai F_{tabel} dengan rumus : $F = F(\frac{dk = n-1}{dk = n-1})$

$$F = F(\frac{dk = 34-1}{dk = 34-1}) = F(33/33)$$

Karena tidak terdapat dalam tabel F maka dicari dengan interpolasi sebagai berikut :

$$\left.\begin{matrix} F_{(0,05)(30/32)} = 1{,}82 \\ F_{(0,05)(40/32)} = 1{,}76 \end{matrix}\right\} F_{(0,05)(33/32)} = 1{,}82 - \frac{9}{10}(0{,}06) = 1{,}766$$

$$\left.\begin{matrix} F_{(0,05)(30/40)} = 2{,}20 \\ F_{(0,05)(40/40)} = 2{,}11 \end{matrix}\right\} F_{(0,05)(39/40)} = 2{,}20 - \frac{9}{10}(0{,}09) = 2{,}119$$

$$\left.\begin{matrix} F_{(0,05)(33/32)} = 1{,}766 \\ F_{(0,05)(32/40)} = 2{,}119 \end{matrix}\right\} F_{(0,05)(33/33)} = 1{,}766 - \frac{1}{2}(-0{,}353) = 1{,}943$$

3. Kriteria Uji : Jika $F_{hitung} < F_{tabel}$ maka homogen
   Karena nilai $F_{hitung} = 2{,}174 \geq F_{tabel} = 1{,}943$ maka tidak homogen

Karena data pretest berdistribusi normal dua-duanya dan tidak homogen maka dilanjutkan keUji t'.

Uji t' (Uji Kesamaan Dua Rata-rata)

1. Merumuskan hipotesis nol dan hipotesis alternatifnya

   Ho : Tidak terdapat perbedaan kemampuan awal yang signifikan antara kelas eksperimen dan kelas kontrol

   Ha : Terdapat perbedaan kemampuan awal yang signifikan antara kelas eksperimen dan kelas kontrol

2. Menentukan nilait 't$_{hitung}$ dengan rumus : $t'_{hitung} = \dfrac{\overline{\chi_1} - \overline{\chi_2}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

$$t'_{hitung} = \frac{11{,}56 - 10{,}09}{\sqrt{\dfrac{(4{,}998)^2}{34} + \dfrac{(3{,}39)^2}{34}}} = 1{,}24$$

3. Menentukan kriteria pengujian hipotesis: Ho diterima jika:

$$-\frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} < t' < \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

Dengan: $w_1 = \dfrac{s_1^2}{n_1}$; $w_2 = \dfrac{s_2^2}{n_2}$; $t_1 = t_\alpha(n_1\text{-}1)$; $t_2 = t_\alpha(n_2\text{-}1)$

$$w_1 = \frac{(4{,}99)^2}{34} = 0{,}732 \qquad w_2 = \frac{(3{,}39)^2}{34} = 0{,}338$$

$$t_1 = t_{0.05}(34-1) = t_{0.05}(33) = 2{,}04 - (0{,}5)(2{,}04 - 2{,}02) = 2{,}03$$

$$t_2 = t_{0.05}(34-1) = t_{0.05}(33) = 2{,}04 - (0{,}4)(2{,}04 - 2{,}02) = 2{,}032$$

maka

$$-\frac{(0{,}732)(2{,}03) + (0{,}338)(2{,}032)}{(0{,}732) + (0{,}338)} < 1{,}24 < \frac{(0{,}732)(2{,}03) + (0{,}338)(2{,}032)}{(0{,}732) + (0{,}338)}$$

$-3{,}061 < 1{,}24 < 3{,}061$

Karena nilai t' = 1,24 berada pada daerah penerimaan Ho, maka dapat disimpulkan bahwa: "Tidak terdapat perbedaan kemampuan awal yang signifikan antara kelas eksperimen dengan kelas kontrol".

7.3. Soal-soal Latihan

1. Periksa, apakah kedua varians dari data kelompok X dan Y berikut homogen?

Kode Siswa	X	Kode Siswa	Y
E-1	35	K-1	35
E-2	25	K-2	30
E-3	35	K-3	45
E-4	15	K-4	25
E-5	10	K-5	15
E-6	15	K-6	35
E-7	35	K-7	30
E-8	5	K-8	40
E-9	15	K-9	25
E-10	5	K-10	25
E-11	45	K-11	5
E-12	35	K-12	15
E-13	55	K-13	25
E-14	10	K-14	15
E-15	35	K-15	45
E-16	20		
Banyak Data	n_x = …….		n_y = …….
Nilai Rata-rata	$\bar{x}$ = …….		$\bar{y}$ = …….
Simpangan Baku	S_x = …….		S_y = …….
Varians	V_x = …….		V_y = …….

2. Seorang guru matematika menetapkan standar kelulusan sebesar 70. Diambil sampel random 10 siswa dan diperiksa tes matematikanya. Dari hasil pemeriksaan masing-masing siswa diperoleh nilai: 75 60 68 78 81 72 62 61 72 65. Benarkah nilai matematika siswa lebih tinggi dari standard yang ditetapkan? (gunakan $\alpha = 0{,}01$).
3. Kepada 10 siswa diberikan pelajaran tambahan Bahasa Indonesia. Jika hasil tes ke-10 siswa tersebut seperti di bawah ini, apakah ada perbedaan hasil belajar Bahasa Indonesia antara sebelum dan sesudah pelajaran tambahan? (gunakan $\alpha = 0{,}05$).

Subyek	1	2	3	4	5	6	7	8	9	10

Sebelum	48	50	63	77	64	60	55	62	60	40
Sesudah	60	55	62	70	70	75	50	60	68	60

4. Misal kita ingin mengetahui perbandingan "Hasil belajar siswa dalam pembelajaran Bahasa Indonesia antara yang mendapatkan pembelajaran *X* dengan siswa yang mendapatkan pembelajaran Y". Berdasarkan hasil Post-tes terhadap dua kelas X dan Y diperoleh data sebagai berikut.

X :	50	50	55	55	65	65	65	65
	70	70	70	70	70	75	75	80
	80	80	85	85	90	90		
Y :	40	60	60	65	65	65	70	70
	70	70	75	75	80	80	80	80
	85	85	85	90	90	90		

Dengan menggunakan α = 0,05 dan melalui uji dua pihak, periksa apakah "Hasil belajar siswa dalam pembelajaran Bahasa Indonesia antara yang mendapatkan pembelajaran *X* dengan siswa yang mendapatkan pembelajaran Y" berbeda?

DAFTAR PUSTAKA

Arifin Zainal. 2012. *Evaluasi Pembelajaran.* Jakarta: Direktorat Jenderal Bain, L. J. (1992). *Introduction to Probability and Mathematical Statistics.* Second Edition. United State of America: Advition of Wadworth, INC.

Boediono dan koster, W. (2001). *Teori dan Aplikasi Statistika dan Probabilitas*. Bandung : PT. Remaja Rosdakarya

Cononer, W. J. (1999). *Practical Nonparametric Statiustic*. Thirt Edition. Texas, America : John Wiley & Sons, INC.

Darmawi, A. (*Statistik Parametrik Menggunakan Sofwere IBM SPSS 22***.** Bahan pengantar kuliah Mahasiswa transfer D3 ke S1 Keperawatan. https://www.academia.edu/35761654/Statistik_Parametrik_Menggunakan_Sofwere_IBM_SPSS_22_dan_Interpretasi_Hasil_Keluarannya. Tersedia [5-6-2023]

Daniel, Wayne, W. (1989). *Statistik Nonparametrik Terapan*. Jakarta: PT. Gramedia.

Hair JF, Black WC, Babin BC, Anderson RE & Tatham RL. (2006). *Multivariate Analysis*. Pearson, Prentice Hall

Kuswardi dan Mutiara, E. (2004). Statistik Berbasis Komputer untuk Orang-orang Nonstatistik. Jakarta : PT Alex Media Komputindo.

Meir, KJ & Brudney JL. (1987). *Applied Statistics for Public Administration*. Brooks/Cole.

Mendenhall, W. (1993). *Beginning Statistics: A to Z*. Duxbury Press

Nugana, E. (1993) *Statistik Untuk Penelitian*. Bandung: Permadi

Ruseffendi E.T. (1994). *Dasar-Dasar Penelitian Pendidikan dan Bidang Non-Eksakta Lainnya.* Semarang: IKIP Semarang Press.

Ruseffendi E.T. (1998). *Statistika Dasar untuk Penelitian Pendidikan.* Bandung: IKIP Bandung Press.

Santoso, S. (2003). *Mengatasi Berbagai Masalah Statistik dengan SPSS versi II.5*. Jakarta: PT Alex Media Komputindo

Santoso, S. (2005). Menggunakan SPSS untuk Statistik Non Parametrik. Jakarta: PT Alex Media Komputindo.

Siegel, S. 1997. *Statistik Nonparametrik untuk Ilmu-ilmu Sosial*. Jakarta: Gramedia.

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

StatSoft, Inc. (2006). Elementary Concept. *Electronic Statistics Textbook*. Tulsa, OK: StatSoft. WEB: http://www.statsoft.com/textbook/stathome.html

StatSoft, Inc. (2006). Basic Concepts. *Electronic Statistics Textbook*. Tulsa, OK: StatSoft. WEB: http://www.statsoft.com/textbook/stathome.html

Sudjana. (1996). *Metoda Statistika*. Bandung: Tarsito

Sugiyono. (2003). *Statistik Nonparametrik untuk Penelitian*. Bandung: CV. Alfabeta.

Sugiyono. (2003). *Statistika untuk Penelitian*. Bandung: CV Alfabeta.

Wonnacott TH & Wonnacott RJ. (1990). *Introductory Statistics for Business and Economics.* Willey

Zanten, V. W. (1980). *Statistik untuk Ilmu- ilmu Sosial*. Jakarta: Gramedia.

Zhafran Ghani Al Rafisqy. 2023. *Skala Pengukuran dalam Ilmu Statistik*. https://ekspektasia.com/skala-pengukuran/

BAB 8

UJI KOMPARATIF SKALA KATEGORIK (KORELASI DAN ASOSIASI)

Oleh Romi Mesra

8.1. Pendahuluan

Banyak penelitian yang bersifat kuantitatif (Sugiyono, 2019) menggunakan variabel kualitatif, baik dalam ilmu biomedis maupun ilmu sosial. Variabel-variabel ini juga dikenal sebagai kategori dan besarnya dinyatakan dalam frekuensi kemunculan masing-masing kategorinya. Variabel kualitatif dapat dibagi lagi menjadi dikotomis (misalnya, jenis kelamin, kematian, penyembuhan), ordinal (misalnya, stadium kanker, amplitudo denyut, kelas fungsional, fototipe, risiko anestesi) atau politomus multinomial (misalnya, orientasi seksual, tipe ABO , status perkawinan, agama, ras, jenis aneurisma, jenis ulkus kronis).

Menurut statistik frequentist, probabilitas proporsi peristiwa yang dipilih secara acak, tanpa penggantian kasus, dapat digeneralisasikan dari distribusi chi-kuadrat, sementara uji chi-kuadrat Pearson didasarkan pada perbedaan antara frekuensi yang diamati dan frekuensi yang secara ideal diharapkan untuk setiap kategori dan dapat digunakan untuk membandingkan seberapa cocok suatu sampel dengan distribusi yang diketahui (misalnya, untuk perbandingan dengan literatur) atau independensi antara sampel yang berbeda.10 Meskipun popularitas uji chi-square Pearson, metode lain seperti uji G

(rasio kemungkinan) dan uji Goodman (kontras antar proporsi) juga digunakan untuk membandingkan proporsi. Namun, keunggulan absolut di antara mereka belum ditentukan secara sistematis (Suharsaputra, 2012).

Korelasi adalah ukuran kekuatan hubungan antara dua variabel. Korelasi tidak menunjukkan kausalitas dan tidak digunakan untuk membuat prediksi; alih-alih mereka membantu mengidentifikasi seberapa kuat dan ke arah mana dua variabel berkovariasi dalam suatu lingkungan. Dalam konteks pengembangan kriteria nutrisi, analisis korelasi adalah alat yang ampuh untuk mengeksplorasi variabel mana yang mungkin terkait kuat dengan konsentrasi nutrisi (Sugiyono, 2016).

Jenis (Sarwono, 2006):

• Pearson (parametrik, mengasumsikan hubungan linier)

• Spearman (non-parametrik, bisa non-linear)

• Kendall's Tau (non-paramentris, bisa non-linear) (R. Lyman-Ott, 2010).

8.2. Uji Komparatif Skala Kategorik (Korelasi)

Korelasi mengkuantifikasi sejauh mana dua variabel kuantitatif, X dan Y, "berjalan bersama". Ketika nilai X yang tinggi dikaitkan dengan nilai Y yang tinggi, ada korelasi positif. Ketika nilai X yang tinggi dikaitkan dengan nilai Y yang rendah, ada korelasi negatif. Kumpulan data ilustrasi. Kami menggunakan kumpulan data bicycle.sav untuk mengilustrasikan metode korelasional. Dalam set data cross-sectional ini, setiap observasi mewakili sebuah lingkungan. Variabel X adalah status sosial ekonomi yang diukur sebagai persentase anak-anak di lingkungan yang menerima makan siang gratis atau potongan biaya di sekolah. Variabel Y adalah penggunaan helm sepeda yang diukur sebagai persentase pengendara sepeda di lingkungan yang memakai helm. Dua belas lingkungan dianggap (Yusuf, 2016).

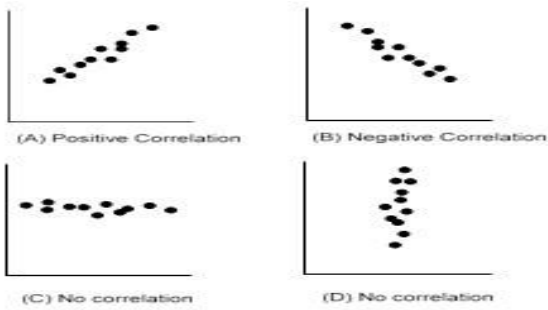
Tabel 8.1 Set Data Cross-Sectional

Neighborhood	X (% receiving reduced-fee lunch)	Y (% wearing bicycle helmets)
Fair Oaks	50	22.1
Strandwood	11	35.9
Walnut Acres	2	57.9
Discov. Bay	19	22.2
Belshaw	26	42.4
Kennedy	73	5.8
Cassell	81	3.6
Miner	51	21.4
Sedgewick	11	55.2
Sakamoto	2	33.3
Toyon	19	32.4
Lietz	25	38.4

(Sumber:Yusuf, 2016)

Tiga adalah dua belas pengamatan (n = 12). Secara keseluruhan, = 30,83 dan = 30,883. Kami ingin mengeksplorasi hubungan antara status sosial ekonomi dan penggunaan helm sepeda. Perlu dicatat bahwa outlier (84, 46.6) telah dihapus dari kumpulan data ini sehingga kita dapat mengukur hubungan linier antara X dan Y.

Scatter Plot



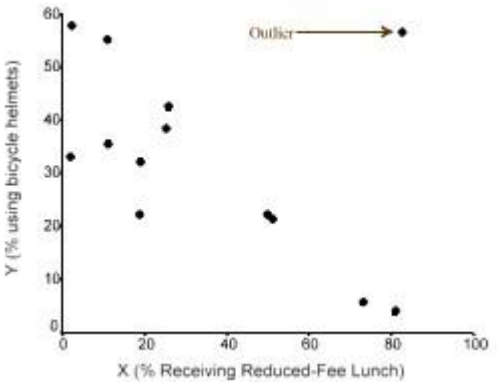
Gambar 8.1 Scatter Plot
(Sumber: Yusuf, 2016)

Langkah pertama adalah membuat sebaran data. “Tidak ada alasan untuk gagal merencanakan dan melihat.”1

Secara umum, Scatter Plot dapat mengungkapkan:

1. Korelasi positif (nilai X tinggi terkait dengan nilai Y tinggi)
2. Korelasi negatif (nilai X tinggi terkait dengan nilai Y rendah)
3. Tidak ada korelasi (nilai X sama sekali tidak memprediksi nilai Y).

Pola-pola ini ditunjukkan pada gambar di bawah berikut.



Gambar 8.2 Pola Scatter Plot
(Sumber: Yusuf, 2016)

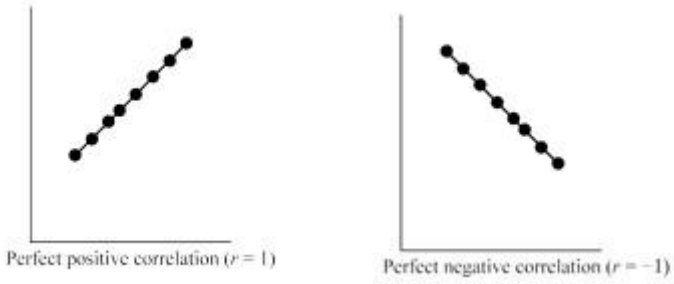
Contoh ilustrasi. Scatter Plot dari data ilustratif ditampilkan di sebelah kanan. Plot mengungkapkan bahwa nilai X yang tinggi dikaitkan dengan nilai Y yang rendah. Artinya, sebagai jumlah anak yang menerima pengurangan biaya makan di sekolah meningkat, tingkat penggunaan helm sepeda menurun' ada korelasi negatif.

Selain itu, terdapat pengamatan yang menyimpang ("outlier") di kuadran kanan atas. Penyimpangan tidak boleh diabaikan penting untuk dikatakan

sesuatu tentang pengamatan yang menyimpang.2 Apa yang harus dikatakan sebenarnya tergantung pada apa yang bisa dipelajari dan apa yang diketahui. Ada kemungkinan pembelajaran dari outlier lebih penting daripada objek utama penelitian. Dalam data

ilustratif, misalnya, kami memiliki sekolah SES rendah dengan catatan keamanan yang buruk. Apa yang memberi?

Koefisien Korelasi



Gambar 8.3 Koefisien Korelasi
(Sumber: Yusuf, 2016)

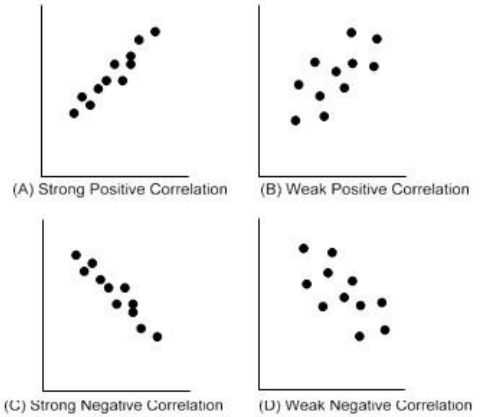
Gagasan Umum

Koefisien korelasi (dilambangkan r) adalah statistik yang mengukur hubungan antara X dan Y dalam suku-suku bebas unit. Ketika semua titik dari plot pencar jatuh langsung pada garis dengan kemiringan ke atas, r = +1; Ketika semua titik jatuh

lurus pada bidang miring ke bawah, r = –1.

Korelasi sempurna seperti itu jarang ditemui. Kita masih perlu mengukur kekuatan korelasional, –didefinisikan sebagai derajat ke mana titik data mematuhi garis tren imajiner yang melewati "awan pencar". Korelasi yang kuat dikaitkan dengan awan pencar yang melekat erat pada garis tren imajiner. Korelasi yang lemah dikaitkan dengan awan pencar yang sedikit mengikuti garis tren.

Semakin dekat r ke +1, semakin kuat korelasi positifnya. Semakin dekat r ke –1, semakin kuat korelasi negatifnya. Contoh korelasi kuat dan lemah ditunjukkan di bawah ini. Catatan: Kekuatan korelasional tidak dapat diukur secara visual. Itu terlalu subyektif dan mudah dipengaruhi oleh penskalaan sumbu. Mata bukanlah penilai kekuatan korelasional yang baik.



Gambar 8.4 Korelasi Kuat dan Lemah
(Sumber: Yusuf, 2016)

Koefisien Korelasi Pearson

Untuk menghitung koefisien korelasi, biasanya Anda memerlukan tiga jumlah kuadrat (SS) yang berbeda. Jumlah kuadrat variabel X, jumlah kuadrat variabel Y, dan jumlah perkalian silang XY.

Jumlah kuadrat untuk variabel X adalah:

(1) $$SS_{XX} = \sum (x_i - \bar{x})^2$$

Statistik ini melacak penyebaran variabel X. Untuk data ilustrasi, = 30,83 dan SS = (50–30,83)2 + (11–30.83)2 + . . . + (25–30,83)2 = 7855,67. Karena statistik ini adalah pembilang varian dari X (s2), maka bisa juga dihitung sebagai SS = (s2)(n–1). Jadi SS = (714.152)(12–1) = 7855.67.

Jumlah kuadrat untuk variabel Y adalah:

(2) $$SS_{YY} = \sum (y_i - \bar{y})^2$$

Statistik ini melacak penyebaran variabel Y dan merupakan pembilang varians Y (s2). Untuk data ilustrasi = 30.883 dan SS = (22.1–30.883)2 + (35.9–30.883)2 + . . . + (38,4–30,883)2 = 3159,68. Sebuah cara alternatif untuk menghitung jumlah kuadrat untuk variabel Y adalah SS = (s2)(n–1). Jadi SS = (287.243)(12–1) =

3159.68.

Akhirnya, jumlah produk silang (SSXY) adalah:

(3) $$SS_{XY} = \sum (x_i - \bar{x})(y_i - \bar{y})$$

Untuk data ilustrasi, SSXY = (50–30.83)(22.1–30.883) + (11–30.83)(35.9–30.883) + . . . + (25–30,83)(38,4 –30,883) = –4231,1333. Statistik ini analog dengan jumlah kuadrat lainnya kecuali digunakan untuk menghitung sejauh mana kedua variabel "berjalan bersama".

Koefisien korelasi (r) adalah,

(4) $$r = \frac{SS_{XY}}{\sqrt{(SS_{XX})(SS_{YY})}}$$

Untuk data ilustrasi, r = $\frac{-42311333}{\sqrt{(7855.67)(3159.68)}} = -0.849.$

Interpretasi Koefisien Korelasi Pearson

Tanda koefisien korelasi menentukan apakah korelasi itu positif atau negatif. Besarnya koefisien korelasi menentukan kekuatan korelasi. Meskipun tidak ada aturan keras dan cepat untuk menjelaskan kekuatan korelasional, saya [dengan ragu-ragu] menawarkan pedoman ini:

Tabel 8.2 Interprestasi Koefisien Korelasi Pearson

$0 < \|r\| < .3$	weak correlation
$.3 < \|r\| < .7$	moderate correlation
$\|r\| > 0.7$	strong correlation

(Sumber: Yusuf, 2016)

Misalnya, $r = -0.849$ menunjukkan korelasi negatif yang kuat.

SPSS: Untuk menghitung koefisien korelasi klik Analyze > Correlate > Bivariate. Kemudian pilih variabel untuk analisis. Beberapa koefisien korelasi bivariat dapat dihitung secara bersamaan dan ditampilkan sebagai matriks korelasi. Mengklik tombol Opsi dan mencentang "Penyimpangan dan kovarian produk silang" akan menghitung jumlah kuadrat.

Koefisien Determinasi

Koefisien determinasi adalah kuadrat dari koefisien korelasi (r2). Untuk data ilustrasi, $r^2 = -0.849^2 = 0.72$. Statistik ini mengkuantifikasi proporsi varian dari satu variabel yang "dijelaskan" (dalam pengertian statistik, bukan pengertian kausal) oleh variabel lainnya. Koefisien determinasi ilustratif sebesar 0,72 menunjukkan bahwa 72% variabilitas penggunaan helm dijelaskan oleh status sosial ekonomi (Gerstman, 2003).

8.3. Uji Komparatif Skala Kategorik (Asosiasi)

Pemilihan jenis uji dilakukan dengan memperhatikan jenis data yang digunakan, jumlah sampel maupun hubungan antar variabel. Secara garis besar pemilihan uji statistik dapat disederhanakan menjadi berikut (Kusumastuti, Khoiron and Achmadi, 2020):

Tabel 8.3 Pengelompokan Uji Statistik

Tabel 1. Pengelompokan uji statistik	JENIS ASOSIASI
SKALA PENGUKURAN	
KOMPARATIF	KORELASI
Tidak berpasangan	berpasangan

	2 kelompok	>2 kelompok	2 kelompok	>2 kelompok	
Kategorik	*Uji t-Independent*	*One-way ANOVA*	*Uji t-dependent*	*Two-way ANOVA*	*Pearson*
Nominal	*Mann Whitney*	*Kruskall Wallis*	*Friedman*		*Sperman*
Ordinal kategorik	*Chi-Squre, Fisher Kolmogoroc-smirnov*		*Wilxocon McNemar, Cohran, Marginal homogeneity wilxocon, friedman*		*Koefisien konringensi Lambda*

(Sumber: Ramadhan dan Sugiyono, 2015)

Kadang akan muncul beberapa pertanyaan sebagai berikut :

1. Apa persamaan dan perbedaan uji korelasi koefisien kontingensi dengan lamda? Kedua uji tersebut digunakan untuk menguji korelasi dua variabel dimana salah satu variabelnya adalah variabel nominal. Perbedaan: Uji korelasi koefisien kontingensi digunakan untuk menguji korelasi antara dua variabel setara sedangkan uji korelasi Lambda untuk variabel yang idak setara.
2. Apa persamaan dan perbedaan uji korelasi spearman dengan uji korelasi Gamma dan Somers'd Keduanya digunakan untuk uji korelasi antara variabel ordinal dengan ordinal Perbedaan: Uji spearman digunakan juga untuk uji korelasi antara variabel numerik dengan ordinal. Uji spearman digunakan juga sebagai alternatif uji pearson, jika syarat uji pearson tidak terpenuhi. Uji korelasi Gamma dan Somers'd digunakan untuk uji korelasi variabel ordinal dengan ordinal dimana kategori variabel ordinal tersebut "sedikit" sehingga dapat dibuat suatu tabel silang B x K.
3. Apa perbedaan uji korelasi Gamma dan Somers'd Uji korelasi Gamma digunakan untuk menguji korealsi antara dua variabel yang setara sedangkan uji korelasi Somers'd untuk dua variabel yang tidak setara.
4. Bagaimana intepretasi hasil uji korelasi Intepretasi hasil uji korelasi didasarkan pada nilai p, kekuatan korelasi, serta

arah korelasinya. Untuk lebih jelasnya Intepretasi uji korelasi dapat dilihat pada tabel di bawah ini.

Uji bivariat korelatif data kategorik (uji asosiasi)

Uji korelasi Gamma dan Sommers (tabel hipotesis korelatif ordinal B x K) digunakan untuk mengetahui hubungan antara dua variabel. Jika ingin mengetahui hubungan antara tingkat penilaian responden terhadap kualitas pelayanan perawat (buruk, sedang, baik) dan pelayanan rumah sakit (buruk, sedang, baik). Berikut rumusan pertanyaan penelitian: "Adakah hubungan antara tingkat penilaian pasien terhadap mutu pelayanan keperawatan dengan mutu pelayanan rumah sakit?" Tabel berikut menunjukkan langkah-langkah untuk menjawab pertanyaan di atas:

Tabel 8.4 Disesuaikan dengan panduan tabel uji hipotesis, langkah-langkah menentukan uji hipotesis

Tabel 8.4 Langkah-langkah Menentukan Uji Hipotesis

No.	Langkah	Jawaban
1.	Tentukan variabel yang dihubungkan	Variabel yang dihubungkan adalah mutu layanan keperawatan (kategorik ordinal) dengan mutu pelayanan rumah sakit (kategorik ordinal)
2.	Tentukan jenis hipotesis	Korelatif
3.	Tentukan masalah skala variabel	Kategorik ordinal
Kesimpulan : Terdapat tiga pilihan uji korelasi antara lain korelasi spearman, Gamma, dan Sommers'd. Anda akan memilih untuk melakukan uji korealsi Gamma dan Sommers'd karena yang akan diuji disini adalah korelasi antara variabel ordinal yang penyajiannya dalam bentuk silang 3 x 3		

(Sumber: Amal., et al, 2018)

Langkah selanjutnya adalah membuka file Gamma; namun, sebelum melakukannya, periksa View Variables yang

telah Anda buat. Berikut langkah selanjutnya: Investigasi Descriptive Statistics Crosstab Isi titik-titik dengan menggunakan Variabel P3 (Kualitas Pelayanan Rumah Sakit). Isilah titik-titik dengan menggunakan Variabel P4 (Layanan Keperawatan). Aktifkan kotak Statistik. Pilih Gamma dan Sommers'd Prosedur telah selesai.

Intepretasi:

1. Output pertama (Crosstab) akan menampilkan tabel silang yang membandingkan kualitas pelayanan keperawatan dengan kualitas pelayanan rumah sakit.
2. Temuan tes Somers disajikan dalam output kedua (Directional Measures). Jika salah satu variabel adalah variabel independen dan yang lainnya adalah variabel dependen, maka digunakan hasil uji Somers'd. Saat mempertimbangkan kualitas layanan rumah, korelasinya adalah 0,028, menunjukkan bahwa tautannya sangat buruk.
3. Keluaran akhir (Ukuran Simetris) menampilkan temuan uji Gamma, yang akan Anda manfaatkan jika lokasi kedua variabel sebanding (tidak ada variabel independen dan dependen).

Uji Gamma menghasilkan nilai korelasi sebesar 0,052, yang menunjukkan bahwa korelasi sebagai variabel independen sangat lemah, maka angka yang Anda gunakan adalah hasil dari uji Somers'd baris kedua.

DAFTAR PUSTAKA

Gerstman, B.B. (2003) '14 : Correlation', *StatPrimer*, 1, pp. 1–9. Available at: http://www.sjsu.edu/faculty/gerstman/StatPrimer/correlation.pdf.

Kusumastuti, A., Khoiron, A.M. and Achmadi, T.A. (2020) *Metode penelitian kuantitatif*. Deepublish.

R. Lyman-Ott (2010) 'Correlation | What Is It? How Does It Work?', pp. 3–4. Available at: http://www.epa.gov/bioindicators/statprimer/index.html.

Sarwono, J. (2006) 'Metode penelitian kuantitatif & kualitatif'.

Sugiyono (2019) *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.

Sugiyono, P.D. (2016) *metode penelitian kuantitatif, kualitatif,dan R&D*, *Alfabeta, cv.*

Suharsaputra, U. (2012) 'Metode penelitian: kuantitatif, kualitatif, dan tindakan'.

Yusuf, A.M. (2016) *Metode penelitian kuantitatif, kualitatif & penelitian gabungan*. Prenada Media.

BAB 9

UJI KORELASI

Oleh Hartina Husain, S.Si., M.Stat.

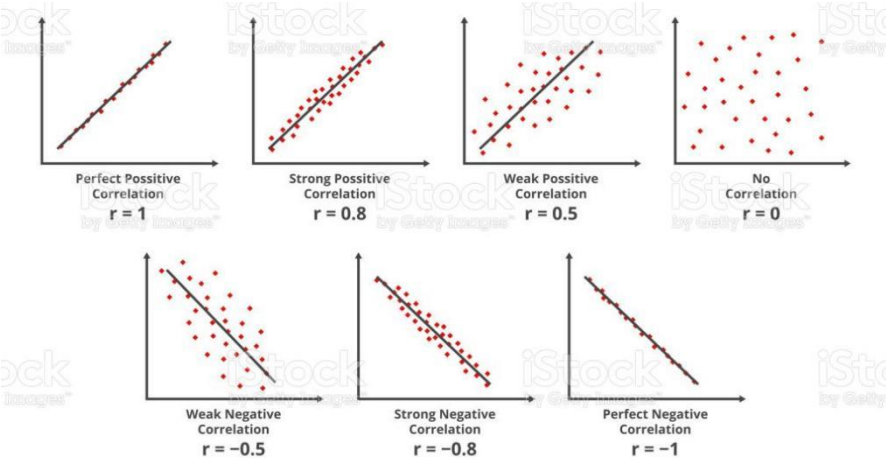
9.1. Pendahuluan

Statistika induktif mulai berkembang pada transisi abad ke -19 menuju abad 20 yang diperkenalkan oleh Karl Pearson. Era ini merupakan awal perkembangan statistika modern yang diterapkan pada berbagai disiplin ilmu pengetahuan. Sebagai contoh pada jurnal biometika yang diterbitkan oleh Karl pearson pada abad 19 merupakan contoh penerapan statistika pada biologi mengenai masalah hereditas dan evolusi. Tidak hanya itu, Pearson juga terhitung sudah memiliki 18 *paper* hingga tahun 1912 yang membahas kontribusi matematika pada teori evolusi berbasis analisis regresi dan korelasi (Sunaryo *et al.*, 2003).

Analisis korelasi merupakan suatu alat statistik yang digunakan untuk mengetahui seberapa besar hubungan antara variabel - variabel penelitian. Adapun variabel yang dimaksud ini mengacu pada karakteristik dari objek penelitian yang terdiri atas variabel dependen (Y) dan variabel independen (X). Variabel dependen (Y) atau biasa disebut variabel respon merupakan variabel yang dipengaruhi, sedangkan variabel independen (X) yaitu variabel yang mempengaruhi (Sulistiyowati and Astuti, 2017). Misalnya dilakukan penelitian pada sekelompok mahasiswa untuk mengetahui adanya hubungan antara motivasi belajar dan Indeks Prestasi (IP) mahasiswa. Dari contoh tersebut dapat ditentukan bahwa variabel independennya yaitu motivasi belajar dan variabel dependennya yaitu indeks prestasi.

Nilai koefisien korelasi yang diperoleh mencerminkan keeratan hubungan dan arah hubungannya. Nilainya berkisar dari nol hingga satu di mana semakin mendekati satu maka semakin erat hubungannya. Begitu pun sebaliknya yaitu semakin

mendekati nol maka hubungan antara variabel sangat kecil bahkan dapat diabaikan. Sementara itu arah hubungan terlihat dari tanda positif (+) atau negatif (-) pada koefisien korelasi. Nilai positif menandakan hubungan antara variabel searah yaitu semakin besar nilai variabel independennya maka semakin tinggi pula nilai variabel dependennya, pun sebaliknya. Namun nilai negatif menandakan bahwa hubungannya timbal balik atau tidak searah. Artinya semakin besar nilai variabel independennya maka semakin kecil nilai variabel dependennya, begitu pun sebaliknya (Budiwanto, 2017).



Gambar 9.1 Bentuk-bentuk grafik pencar

(Sumber: istockphoto.com)

Gambar 9.1 menunjukkan bagaimana korelasi data dilihat dari pola grafik pencar data. Semakin besar korelasi yang dimiliki, maka titik-titik pasangan data cenderung mendekati garis lurus. Korelasi sempurna terjadi jika titik-titik data tepat berada pada garis tersebut. Sementara itu, pada grafik pencar yang tidak dapat dinyatakan garis yang mewakili maka dikatakan tidak ada korelasi. Terlihat pula pergerakan data yang menunjukkan arah hubungannya di mana garis yang mengarah ke atas menunjukkan korelasi positif. Hal ini karena adanya peningkatan nilai X

menyebabkan peningkatan juga pada nilai Y, ataupun sebaliknya. Namun, garis yang mengarah ke bawah menunjukkan korelasi negatif yang berarti peningkatan nilai X menyebabkan penurunan nilai Y.

Nilai r umumnya digunakan sebagai simbol dari koefisien korelasi yang menunjukkan hubungan linear antara variabel. Namun perlu diperhatikan bahwa nilai r ini tidak menunjukkan hubungan sebab-akibat dari variabel melainkan didasarkan pada konteks spesifik dari data yang dianalisis. Sebagai bahan acuan, dapat digunakan Tabel 11.1 dalam membuat interpretasi nilai korelasi (Supriadi, 2019).

Tabel 9.1 Interpretasi nilai r

Nilai r	Interpretasi
$\mathbf{0 \leq r < 0,2}$	Sangat rendah
$\mathbf{0,2 \leq r < 0,4}$	Rendah
$\mathbf{0,4 \leq r < 0,6}$	Cukup kuat
$\mathbf{0,6 \leq r < 0,8}$	kuat
$\mathbf{0,8 \leq r \leq 1}$	Sangat kuat

(Sumber: Penulis, 2023. Diolah)

Uji korelasi dapat dibedakan menjadi beberapa jenis bergantung pada jenis data yang dianalisis dan skala pengukuran datanya. Pada bab ini dijelaskan 4 jenis korelasi yaitu korelasi Product Moment Pearson, korelasi Spearman, korelasi Point-Biserial, dan korelasi Phi.

9.2. Korelasi Product Moment Pearson

Koerelasi Product Moment Pearson merupakan analisis yang digunakan untuk mengetahui hubungan/korelasi antara dua variabel yang masing-masing memiliki skala pengukuran interval dan atau berskala rasio (Rosalina, 2023). Namun, perlu

diperhatikan asumsi normalitas dan linearitas data harus terpenuhi untuk menggunakan analisis ini.

Misalkan kedua variabel yang diukur dilambangkan dengan X dan Y maka nilai koefisien korelasi diperoleh menggunakan rumus

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

Atau

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

Dimana

r_{XY} : Koefisien korelasi

X : Variabel independen

Y : Variabel dependen

n : Banyak sampel yang diteliti

Contoh :

Sebuah perusahaan *smartphone* ingin mengetahui apakah terdapat hubungan antara jumlah iklan yang ditempatkan di berbagai *platform* media sosial dan volume penjualan produknya. Dengan demikian, dilakukan pengumpulan data iklan yang telah diterbitkan dan jumlah penjualan bulanan Tahun 2022 untuk mengevaluasi apakah iklan memiliki dampak yang signifikan terhadap penjualan.

Tabel 9.2 Jumlah iklan dan volume penjualan perusahaan smartphone

Bulan	Jumlah Iklan (X)	Volume Penjualan (Y)

Januari	18	35
Februari	20	47
Maret	16	19
April	21	46
Mei	16	16
Juni	17	44
Juli	16	28
Agustus	`15	36
September	18	44
Oktober	13	25
November	17	49
Desember	13	18

(Sumber: Penulis, 2023. Diolah)

Dari Tabel 9.2 didapatkan bahwa variabel yang mempengaruhi/variabel independen yaitu jumlah iklan sehingga dapat disimbolkan dengan X. Sementara itu, variabel yang dipengaruhi/variabel dependen yaitu volume penjualan yang disimbolkan dengan Y. Selanjutnya, untuk menentukan koefisien korelasi dapat digunakan tabel penolong seperti pada Tabel 9..3.

Tabel 9.3 Tabel penolong menghitung koefisien korelasi

X	Y	X^2	Y^2	XY
18	35	324	1225	630
20	47	400	2209	940

16	19	256	361	304
21	46	441	2116	966
16	16	256	256	256
17	44	289	1936	748
16	28	256	784	448
15	36	225	1296	540
18	44	324	1936	792
13	25	169	625	325
17	49	289	2401	833
13	18	169	324	234
$\sum \boldsymbol{X}$=200	$\sum Y$=407	$\sum X^2$ =3398	$\sum Y^2$=15469	$\sum XY$=7016

(Sumber: Penulis, 2023. Diolah)

$$r_{XY} = \frac{12(7016) - (200)(407)}{\sqrt{(12(3398) - 200^2)(12(15469) - 407^2)}}$$

$$= \frac{84192 - 81400}{\sqrt{(40776 - 40000)(185628 - 165649)}}$$

$$= \frac{2792}{\sqrt{(776)(19979)}}$$

$$= \frac{2792}{3937{,}474} = 0{,}71$$

Hasil perhitungan diperoleh r = 0,71 yang menunjukkan korelasi antara jumlah iklan yang ditempatkan diberbagai *platform* media sosial dan volume penjualan produk mengindikasikan adanya hubungan yang kuat antara kedua

variabel tersebut. Adapun tanda koefisien korelasi positif sehingga menunjukkan arah hubungannya searah yang diartikan bahwa semakin besar jumlah iklan maka semakin besar pula volume penjualan produk, begitu pun sebaliknya.

9.3. Korelasi Spearman

Korelasi Spearman merupakan metode statistik yang digunakan untuk mengukur hubungan antara dua variabel yang memiliki skala ordinal/peringkat. Berbeda dengan korelasi product moment pearson yang mengukur hubungan linear antara variabel-variabel yang berkelanjutan, korelasi spearman bekerja untuk memberi peringkat data pada kedua variabel.

Metode ini digunakan pada data yang tidak memenuhi asumsi distribusi normal. Langkah-langkah analisis korelasi spearman yaitu sebagai berikut (Wahyudi, 2010):

1. Jika data berupa skala interval atau rasio maka dilakukan transformasi data ke skala ordinal. Variabel X dan variabel Y masing-masing diberikan peringkat baik dimulai dari nilai yang paling kecil ataupun sebaliknya. Apabila terdapat data yang bernilai sama, maka peringkat didasarkan pada rata-rata dari peringkat-peringkat data tersebut.
2. Mencari selisih peringkat antara variabel X dan variabel Y $(d = rank\ (X) - rank\ (Y))$.
3. Menghitung nilai koefisien korelasi spearman (r_s) menggunakan rumus berikut :

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

Contoh :

Suatu penelitian dilakukan untuk mengetahui hubungan antara jumlah jam belajar dan nilai ujian mahasiswa jurusan Ekonomi di Universitas X. Terdapat 10 mahasiswa yang dijadikan sebagai sampel penelitian dengan data berikut:

Tabel 9.4 Data jumlah jam belajar dan nilai ujian mahasiswa

Mahasiswa	Jumlah jam (X)	Nilai Ujian (Y)
A	3	55
B	5	75
C	6	78
D	10	93
E	4	68
F	6	53
G	1	40
H	9	95
I	8	82
J	7	70

(Sumber: Penulis, 2023. Diolah)

Dengan menggunakan analisis korelasi Spearman maka data Tabel 9.4 diubah menjadi peringkat/skala ordinal.

Tabel 9.5 Tabel perhitungan analisis korelasi spearman

Rank (X)	Rank (Y)	d	d^2
2	3	-1	1
4	6	-2	4
5,5	7	-1,5	2,25
10	9	1	1
3	4	-1	1
5,5	2	3,5	12,25

1	1	0	0
9	10	-1	1
8	8	0	0
7	5	2	4
			$\sum d^2$= 26,5

(Sumber: Penulis, 2023. Diolah)

$$r_s = 1 - \frac{6(26,5)}{10(10^2 - 1)}$$
$$= 1 - \frac{159}{10(99)}$$
$$= 1 - \frac{159}{990}$$
$$= 1 - 0,16 = 0,84$$

Nilai koefisien korelasi sebesar 0,84 menunjukkan bahwa hubungan antara jumlah jam belajar dan hasil belajar mahasiswa sangat kuat dengan arah hubungannya positif. Dengan demikian, semakin besar jumlah jam belajar maka semakin besar pula hasil belajar mahasiswa, begitu pun sebaliknya.

9.4. Korelasi Point-Biserial

Korelasi Point-Biserial adalah metode statistik yang digunakan dalam mengukur hubungan antara variabel biner (dikotomi) dan variabel berskala interval/rasio. Variabel biner yaitu variabel yang hanya memiliki dua kemungkinan nilai seperti benar atau salah, hidup atau mati, sukses atau gagal, pria atau wanita, dan lain sebagainya (Suyanto *et al.*, 2018). Untuk memudahkan dalam perhitungan, variabel biner dapat disimbolkan X dan variabel berskala interval/rasio disimbolkan dengan Y. Adapun rumus yang digunakan dalam menghitung korelasi point biserial yaitu (Brown, 2001) :

$$r_{pb} = \frac{M_P - M_q}{S_y}\sqrt{pq}$$

Dimana

r_{pb}: Korelasi point-biserial

p : Proporsi kelompok data Y bernilai 1

q : Proporsi kelompok data Y bernilai 0 $(q = 1 - p)$

M_P: Rata-rata kelompok data Y bernilai 1

M_q : Rata-rata kelompok data Y bernilai 0

S_y : Standar deviasi data Y

Contoh :

Misalkan dilakukan penelitian untuk mengetahui apakah terdapat hubungan antara jenis kelamin dan tinggi badan dalam satu kelompok mahasiswa.

Tabel 9.6 Data jenis kelamin dan tinggi badan

Mahasiswa	Jenis Kelamin (X)	Tinggi badan (cm) (Y)
A	1	175
B	0	160
C	1	178
D	1	170
E	0	155
F	1	167
G	1	160

H	0	158

(Sumber: Penulis, 2023. Diolah)

Tabel 9.6 menunjukkan bahwa terdapat dua variabel yaitu jenis kelamin yang bernilai dikotomi (1 untuk pria, 0 untuk wanita) dan tinggi badan yang berskala rasio. Kemudian dapat ditentukan:

$$p = \frac{5}{8} = 0{,}625$$

$$q = \frac{3}{8} = 0{,}375$$

$$M_P = \frac{175 + 178 + 170 + 167 + 160}{5} = 170$$

$$M_q = \frac{160 + 155 + 158}{3} = 157{,}67$$

$S_y = 7{,}87$ (menggunakan rumus standar deviasi)

Selanjutnya

$$r_{pb} = \frac{170 - 157{,}67}{7{,}87}\sqrt{(0{,}625)(0{,}375)}$$

$$= \frac{12{,}33}{7{,}87}\sqrt{0{,}23}$$

$$= 1{,}57(0{,}48)$$

$$= 0{,}75$$

Hasil perhitungan korelasi diperoleh sebesar 0,75 yang menunjukkan korelasi positif yang kuat antara jenis kelamin dan tinggi badan. Artinya, pria cenderung memiliki tinggi badan yang lebih tinggi dibandingkan dengan wanita dalam kelompok mahasiswa yang dianalisis.

9.5. Korelasi Phi

Korelasi Phi merupakan metode statistik untuk mengukur hubungan antara dua variabel biner (dikotomi) dalam bentuk tabel kontingensi 2x2. Tabel 11.7 dibawah ini berguna dalam mengorganisasi data menjadi empat sel dimana tiap sel mewakili kombinasi antara dua variabel.

Tabel 9.7 Tabel kontingensi 2x2

Variabel		Y		Total
		1	0	
X	1	a	b	a+b
	0	c	d	c+d
Total		a+c	b+d	n

(Sumber: Penulis, 2023. Diolah)

Dengan rumus korelasi phi yaitu (Syafril, 2019) :

$$r_{phi} = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

Contoh :

Suatu penelitian dilakukan untuk menguji hubungan antara jenis minuman yang dikonsumsi (teh/kopi) dengan kebiasaan merokok (perokok/bukan perokok) pada sekelompok individu.

Tabel 9.8 Data jenis minuman dan kebiasaan merokok

Responden	Jenis Minuman (X) 1 = Teh, 0 = Kopi	Kebiasaan Merokok (Y) 1 = Perokok, 0 = Bukan Perokok
1	1	0

2	0	1
3	1	0
4	1	1
5	0	0
6	1	0
7	0	1
8	0	0

(Sumber: Penulis, 2023. Diolah)

Selanjutnya data Tabel 11.8 diubah menjadi tabel kontingensi 2x2

Tabel 9.9 Tabel kontingensi perhitungan korelasi phi

Variabel		Kebiasaan Merokok		Total
		1 (Perokok)	0 (Bukan perokok)	
Jenis Minuman	1 (Teh)	1	3	4
	0 (Kopi)	2	2	4
Total		3	5	8

(Sumber: Penulis, 2023. Diolah)

Diperoleh

$$r_{phi} = \frac{(1)(2) - (3)(2)}{\sqrt{(4)(4)(3)(5)}}$$

$$= \frac{2 - 6}{\sqrt{240}}$$

$$= \frac{-4}{15,49} = -0,258$$

Hasil perhitungan korelasi sebesar -0,258 yang berarti hubungan antara jenis minuman dan kebiasaan merokok rendah. Adapun arah hubungannya negatif yang menunjukkan bahwa orang yang jenis minumannya teh cenderung bukan perokok. Begitu pun sebaliknya bahwa orang yang jenis minumannya kopi cenderung memiliki kebiasaan merokok (perokok).

DAFTAR PUSTAKA

Brown, J.D. (2001) 'Point Point--biserial correlation coefficients biserial correlation coefficients', *JLT Testing & Evlution SIG Newsletter*, 5(3), pp. 13–17.

Budiwanto, S. (2017) 'Metode Statistika: Untuk Mengolah Data Keolahragaan', *Fakultas Ilmu Keolahragaan Universitas Negeri Malang 2017*, pp. 1–233.

Rosalina, linda; R. oktarina;Rahmiati;Indra S. (2023) *Buku Ajar Statistika*, *MRI Publisher*. Padang.

Sulistiyowati, W. and Astuti, C.C. (2017) 'Statistika Dasar Konsep Dan Aplikasinya', *Nucl. Phys.*, 13(1), pp. 104–116.

Sunaryo, S. *et al.* (2003) 'Sejarah Perkembangan Statistika Dan Aplikasinya', *Forum Statistika Dan Komputasi*, 8(1), pp. 1–7.

Supriadi, G. (2019) *PENELITIAN PENDIDIKAN Metod1.pdf*.

Suyanto *et al.* (2018) *Analisis Data Penelitian Petunjuk Praktis Bagi Mahasiswa Kesehatan Menggunakan SPSS*.

Syafril (2019) *Statistika Pendidikan*. Edisi Pert. Jakarta: Prenadanamedia Group.

Wahyudi, A. (2010) 'Analisis Korelasi Rank Spearman', *Jurnal Metode Kunatitatif*, p. 13.

BIODATA PENULIS

Zul Fadli, S.E., M.A.P.
Dosen Program Studi Ilmu Administrasi Negara
FISIP Universitas Pattimura

Lahir di Ujung Pandang pada tanggal 12 Juni 1988. Merupakan putra kedua dari pasangan Prof. Dr. H. Imran Ismail, M.S dan dr. Hj. Nurhaedah Azis, M.Kes. Menikah pada tahun 2018 dan dikarunia 2 orang anak. Penulis adalah dosen pada Program Studi Ilmu Administrasi Negara FISIP Universitas Pattimura. Menyelesaikan S1 pada Program Studi Akuntansi Jurusan Akuntansi Keuangan Kampus STIE Indonesia Makassar pada tahun 2011, dan S2 pada Program Studi Ilmu Administrasi Konsentrasi Manajemen Sumber Daya Aparatur Kampus STIA-LAN Makassar pada tahun 2018. Telah menulis beberapa buku, melakukan penelitian dan pengabdian kepada masyarakat dan menerbitkan hasilnya pada beberapa jurnal nasional dan internasional.

BIODATA PENULIS

Yanti Murni, S.E., M.M.

Penulis lahir di Kp. Pili tanggal 07 Februari 1986. Penulis adalah dosen tetap pada Program Studi Manajemen Fakultas Ekonomi Universitas Sumatera Barat. Menyelesaikan pendidikan S1 dan S2 pada Jurusan Manajemen. Motivasi penulis dalam membuat book chapter ini adalah dalam rangka proses pembelajaran. Book chapter ini merupakan karya pertama penulis dalam bentuk buku yang ber ISBN.

BIODATA PENULIS

Marsuddin Musa, S.Si., M.Stat.
Dosen Statistika Program Studi Akuntansi
Sekolah Tinggi Ilmu Ekonomi Enam Enam Kendari

Penulis lahir di Gu-Buton tanggal 27 April 1994. Menyelesaikan pendidikan pada Program Studi D3 Statistika Universitas Halu Oleo, kemudian melanjutkan S1 di Jurusan Statistika Institut Sains dan Teknologi Akprind Yogyakarta, dan S2 di Departemen Statistika Institut Teknologi Sepuluh Nopember. Penulis merupakan dosen tetap di Program Studi Akuntansi Sekolah Tinggi Ilmu Ekonomi Enam Enam Kendari dengan bidang keahlian Statistika. Buku ini merupakan salah satu karya dan Insya Allah secara konsisten akan disusul dengan buku-buku berikutnya.

BIODATA PENULIS

Putri Anugrah Cahya Dewi, M.Pd.
Dosen Program Studi Sistem Informasi Akuntansi
Fakultas Teknologi Informasi dan Desain

Penulis lahir di Timor-Timur tanggal 28 Agustus 1994. Penulis adalah dosen tetap pada Program Studi Sistem Informasi Akuntansi Fakultas Teknologi Informasi dan Desain, Universitas Primakara. Menyelesaikan pendidikan S1 dan S2 pada Jurusan Pendidikan Matematika Universitas Pendidikan Ganesha. Jabatan fungsional terakhir adalah Lektor. Aktif menulis dalam beberapa jurnal nasional terakreditasi. Selain sebagai dosen tetap, penulis juga menjabat sebagai Kepala Lembaga Penjaminan Mutu Universitas Primakara.

BIODATA PENULIS

Ketut Queena Fredlina, S.Si., M.Si.
Dosen Program Studi Teknik Informatika
Fakultas Teknologi Informasi dan Desain

Penulis lahir di Denpasar tanggal 31 Oktober 1990. Penulis adalah dosen tetap pada Program Studi Teknik Informatika Fakultas Teknologi Informasi dan Desain, Universitas Primakara. Penulis memperoleh gelar Sarjana (S.Si.) dari Universitas Udayana, pada Jurusan Matematika, kemudian memperoleh gelar Magister (M.Si.) dari Institut Teknologi Bandung (ITB), dengan fokus pada bidang Matematika. Sebagai seorang dosen, penulis telah terlibat dalam pengajaran dan pembimbingan di Jurusan Teknik Informatika di Universitas Primakara. Ia memiliki pengalaman dalam menyampaikan materi yang kompleks dengan cara yang mudah dipahami oleh mahasiswa. Penulis memiliki minat khusus dalam penelitian yang berkaitan dengan penerapan konsep matematika dalam dunia teknologi. Ia telah menghasilkan beberapa karya ilmiah dan penelitian yang berkontribusi pada pemahaman dan perkembangan di bidang ini.

BIODATA PENULIS

Dr. Darham Wahid, S.Pd, M.Pd

Dr. Darham Wahid, S.Pd,. M.Pd. Tempat Tanggal Lahir, Desa Pedukun Kabupaten Bungo, 27 November 1960. Mengenyam Pendidikan D3 Pendidikan Matematika UT 1998, S1 Ilmu Pendidikan YPM Bangko 1996, S2 Administrasi Pendidikan (Konsentrasi Manajemen Sumberdaya Manusia) Pascasarjana IKIP Padang 1999, dan S3 Ilmu Ekonomi (Konsentrasi Manajemen Sumberdaya Manusia) Fakultas Ekonomi dan Bisnis Universitas Jambi tahun 2022. Semasa aktif banyak berpengalaman baik dipemerintahan maupun di swasta. Pengalaman mengajar antara lain di STIA Setih Setio Muara Bungo dari tahun 1999-2006 dengan mata kuliah Statistika, Kewirausahaan, dan mata kuliah bidang manajemen lainnya. Kemudian dari tahun 2008 sampai sekarang menjadi dosen tetap di Universitas Muara Bungo mengajar pada mata kuliah Statistika ekonomi, Metodologi Penelitian Manajemen dan Bisnis, Matematika Ekonomi, dan mata kuliah bidang Manajemen lainnya. Pengalaman penulis dalam organisasi dimulai dari pendirian Universitas Muara Bungo, pernah menjabat sebagai Wakil Rektor III Bidang Kemahasiswaan dan Alumni, Wakil Rektor II Bidang Umum dan Keuangan, dan terakhir Februari 2023 menjabat sebagai Sekretaris Umum Yayasan Pendidikan Muara Bungo.

BIODATA PENULIS

Dr. H. Nanang, M.Pd.
Staf Dosen Jurusan Teknik Sipil
Institut Teknologi Garut

Penulis lahir di Bandung tanggal 1 Juli 1964. Penulis adalah dosen dpk Kopertis Wilayah VII Surabaya pada Program Studi Pendidikan FKIP Universitas Muhammadiyah Jember dari tahun 1990 s.d. 2000 dan dosen LLDIKTI Wilayah IV Jawa Barat dan Banten pada Jurusan Teknik Sipil Fakultas Teknik Sipil dan Arsitektur Institut Teknologi Garut (ITG) dari Tahun 2000 s.d. sekarang. Dari tahun 1986 s.d 1990, penulis bekerja sebagai guru di SMAN 1 Bandung, SMA 55 Asia Afrika Bandung, SMA Karya Agung Bandung, SMA Swadaya Bandung.

Penulis menyelesaikan: S1 Pendidikan Matematika di IKIP Bandung (UPI) Tahun 1989, S2 Pendidikan Matematika di IKIP Surabaya (UNESA) Tahun 1999, dan S3 Pendidikan Matematika di UPI Bandung Tahun 2009. Penulis menekuni bidang Pembelajaran Matematika. Penulis juga sebagai Widiaiswara Karya Ilmiah dan Tim Penilai DUPAK Kepangkatan Guru-guru di Kabupaten Garut dari tahun 2015 s.d. sekarang. Telah beberapa kali memperoleh hibah penelitian dan pengabdian pada masyarakat dari Universitas Muhammadiyah Jember, Institut Teknologi Garut, dan Ristek Dikti Kemendikbud, serta aktif dalam menulis buku ajar dan kegiatan pertemuan ilmiah.

BIODATA PENULIS

Romi Mesra, S.Pd., M.Pd.
Dosen Program Studi Pendidikan Sosiologi
Fakultas Ilmu Sosial dan Hukum
Universitas Negeri Manado

Penulis menaruh perhatian kepada dunia akademis termasuk berkaitan dengan pembahasan pengantar statistik yang merupakan bagian dari materi beberapa mata kuliah yang peneliti ampu.

Tulisan ini menjadi bagian sumbangsih penulis terhadap dunia pendidikan, semoga tulisan ini bermanfaat dan bisa dijadikan referensi ataupun bahan bacaan bagi para akademisi, peneliti, dan masyarakat pada umumnya.

Penulis buku ini adalah dosen PNS di Program Studi Pendidikan Sosiologi, Universitas Negeri Manado yang juga aktif sebagai content creator pada channel youtube: NALURI EDUKASI serta sebagai Editor In Chief JURNAL PARADIGMA: Journal of Sociology Research and Education. Email Penulis: romimesra@unima.ac.id

BIODATA PENULIS

Hartina Husain, S.Si., M.Stat
Dosen Program Studi Tadris Matematika
Fakultas Tarbiyah Institut Agama Islam Negeri Parepare

Penulis lahir pada 23 Mei 1996 di Samata, Sulawesi Selatan. Penulis meraih gelar sarjana dari Universitas Hasanuddin tahun 2017 pada jurusan Statistika. Kemudian penulis melanjutkan studi magister di Surabaya tepatnya di Institut Teknologi Sepuluh Nopember dengan jurusan yang sama. Setelah kelulusannya di tahun 2021, penulis aktif bekerja sebagai Dosen Tadris Matematika di Institut Agama Islam Negeri Parepare. Aktif menulis artikel atau jurnal terkait pemodelan statistika, regresi non parametrik, dan analisis meta. Selain itu, penulis tertarik mempelajari mengenai *data analyst*, *data science* dan *big data*.