DASAR-DASAR STATISTIKA INFERENSI DALAM PENELITIAN

Dra. Perak Samosir, M.Si

Dr. Wilson Rajagukguk, M.Si. M.A

Dra. Ratnawati, M.Si

DASAR-DASAR STATISTIKA INFERENSI DALAM PENELITIAN

Penulis:

Dra. Perak Samosir,M.Si

Dr.Wilson Rajagukguk, S.Si, M.Si

Dra. Ratnawati, M.S

UKI PRESS

Pusat Penerbitan dan Pencetakan

Buku Perguruan Tinggi

Universitas Kristen Indonesia

Jakarta

2022

DASAR-DASAR STATISTIKA INFERENSI DALAM PENELITIAN

Penulis:

Dra. Perak Samosir,M.Si

Dr.Wilson Rajagukguk, S.Si, M.Si

Dra. Ratnawati, M.Si

Editor:

Dr. Indri Jatmoko, S.Si., MM

ISBN: 978-623-6963-93-7

Penerbit: UKI Press
Anggota APPTI
Anggota IKAPI

Redaksi: Jl. Mayjen Sutoyo No.2 Cawang Jakarta 13630
Telp. (021) 8092425

Cetakan I Jakarta: UKI Press, Mei 2021
Hak cipta dilindungi undang-undang.
Dilarang mengutip atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari penerbit.

KATA SAMBUTAN
REKTOR INSTITUT TEKNOLOGI INDONESIA

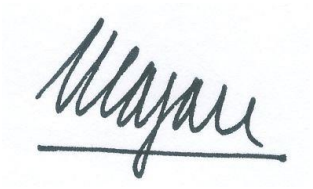
Saya menyambut gembira diterbitkannya buku Dasar-dasar Statistika Inferensi dalam Penelitian yang ditulis oleh Dra. Perak Samosir, M.Si., Dosen Statistik Institut Teknologi Indonesia (ITI), bersama dengan Dr. Wilson Rajagukguk, S.Si., M.Si., M.A., Dosen Universitas Kristen Indonesia dan Dra. Ratnawati, M.Si., Dosen Matematika ITI.

Perlakuan terhadap data, mulai dari pengumpulan, penyajian, analisis dan penafsirannya harus dilakukan dengan sangat cermat dan mengacu pada metode Statistika yang tepat. Dalam lingkungan perguruan tinggi, Statistika Inferensi yang menjadi topik bahasan dalam buku ini sangat bermanfaat sebagai referensi bagi para dosen, peneliti dan mahasiswa dalam menganalisis data yang tersedia. Data yang tersedia ini merupakan sampel yaitu sekumpulan data yang variatif atau tidak unik yang mempunyai induk atau sekumpulan data yang lebih besar jumlahnya atau disebut juga populasi. Statistika Inferensi merupakan metode analisis sampel data tersebut untuk sampai pada hasil, baik berupa peramalan ataupun penarikan kesimpulan. Saat ini penerapannya sangat banyak dilakukan pada domain *Data Science* yaitu *Artificial Intelligence*.

Buku ini disusun berdasarkan pengalaman penulis dalam menulis artikel ilmiah pada jurnal atau prosiding baik nasional, nasional terakreditasi ataupun internasional. Dalam buku ini dibahas secara mendalam dasar-dasar penggunaan Statistika Inferensi dalam penelitian. Publikasi buku ilmiah dalam bentuk buku referensi seperti yang dibuat oleh penulis ini merupakan salah satu upaya untuk memperkaya referensi dalam bidang Statistika, di samping juga dapat menjadi media komunikasi serta penyampaian gagasan pada orang lain sehingga memudahkan komunikasi antar ilmuwan peneliti.

Kepada penulis saya mengucapkan selamat atas keberhasilannya menyelesaikan penulisan buku ini. Semoga buku ini dapat menjadi salah satu referensi bagi para pembaca dalam menganalisis data, sehingga hasilnya valid dan kredibel.

Tangerang Selatan, Agustus 2022.
Rektor Institut Teknologi Indonesia,



(Dr. Ir. Marzan A. Iskandar, IPU)

KATA PENGANTAR

Buku Dasar-dasar Statistika Inferensi dalam Penelitian merupakan penjelasan lengkap tentang hal-hal yang mendasar ketika Statistika Inferensi digunakan dalam suatu penelitian ataupun tulisan ilmiah yang merupakan laporan hasil penelitian. Oleh karenanya buku ini ditujukan kepada pembaca mahasiswa dan peneliti sehingga dapat menggunakan Statistika Inferensi dengan dasar yang baik dan benar.

Saat ini Statistika Inferensi digunakan hampir di semua bidang penelitian. Penelitian-penelitian tersebut umumnya bersifat untuk membuat kesimpulan dari sekumpulan data yang dikumpulkan dengan tujuan tertentu. Bahkan dalam penyelenggaraan urusan-urusan tingkat pemerintahan, kesimpulan dari hasil penelitian digunakan sebagai dasar pengambilan keputusan atau penyusunan kebijakan. Tentu saja proses sampai kepada pengambilan keputusan harus melalui prosedur pengolahan data dan analisis terhadap data yang telah diolah tersebut. Dalam buku ini akan dibahas tentang konsep dasar Statistika dari prosedur tersebut dengan harapan pembaca mahasiswa dan peneliti menguasai teori yang mendasari metode Statistika yang mereka gunakan dalam penelitian.

Penjelasan dalam buku ini dimulai dengan konsep yang paling mendasar tentang landasan kerja Statistika yaitu variasi, reduksi dan generalisasi. Fakta bahwa ketika kita hendak menggunakan sekumpulan data yang disebut sampel; sekumpulan data yang hendak diteliti tersebut bersifat variatif atau tidak unik dan mempunyai induk atau sekumpulan data yang lebih besar yang disebut populasi penelitian. Ketika kita hendak melakukan generalisasi atau kesimpulan terhadap kumpulan data yang lebih besar tersebut, dengan dasar hasil analisis terhadap sekumpulan data yang kita teliti, maka diperlukan dasar teori yang secara ilmiah telah terbukti dapat mentolerir resiko kesalahan dalam pengambilan keputusan tersebut. Dalam hal demikian, Statistika adalah salah satu cabang ilmu yang sudah teruji secara teoritis untuk hal itu. Sebagian besar dari konsep dasar ini menggunakan teori tentang peluang. Dengan demikian pembaca akan dibimbing untuk memahami Statistika dari hal yang paling mendasar yaitu peluang dan distribusi-distribusi peluang yang penting dalam Statistika, khususnya yang digunakan dalam Statistika Inferensi. Kemudian akan dibahas dibahas dasar-dasar teoritis yang menjelaskan bagaimana keterangan dalam sampel dapat digeneralisasikan terhadap keseluruhan populasi penelitian.

Dua pengukuran penting dalam analisis Statistika adalah ukuran pemusatan dan ukuran penyebaran data. Kedua pengukuran ini akan menjadi bagian utama dalam pembahasan Statistika Inferensi khususnya ketika akan melakukan inferensi terhadap populasi berdasarkan sampel acak yang ditarik dari populasi tersebut.

Dalam buku ini juga akan dibahas beberapa contoh penggunaan Statistika Inferensi dalam beberapa artikel ilmiah para penulis yang telah dipublikasikan dalam jurnal atau prosiding baik Nasional ataupun Internasional.

Tentu saja buku ini masih jauh dari sempurna, oleh karenanya penulis berharap saran-saran yang membangun dari pembaca sehingga dapat menyempurnakan isi buku ini.

Akhir kata penulis berharap kiranya buku ini bermanfaat bagi pembaca mahasiswa dan peneliti dalam meningkatkan pemahaman dan kemampuan menggunakan Statistika Inferensi dalam penelitian.

Jakarta, Juli 2022

Penulis

DAFTAR ISI

Kata Pengantar i

Daftar Isi iii

Daftar Tabel vi

Daftar Gambar viii

1 Pendahuluan 1

2 Konsep Peluang 5
- 2.1 Ruang Sampel dan Kejadian 5
- 2.2 Peluang suatu Kejadian 8
- 2.3 Perhitungan Jumlah Titik Sampel 10
- 2.4 Distribusi Peluang dan Grafik 12
- 2.5 Peluang Bersyarat, Aturan Perkalian dan Aturan Penjumlahan 13

3 Distribusi Peluang Teoritis 20
- 3.1 Peubah Acak 20
- 3.2 Distribusi Peluang Peubah Acak Diskrit 21
- 3.3 Distribusi Peluang Peubah Acak Kontinu 22
- 3.4 Ekspektasi (Nilai Harapan atau Harapan Matematik) 22
- 3.5 Beberapa Distribusi Peluang Teoritis 24
  - 3.5.1 Beberapa distribusi peluang peubah acak diskrit 25
  - 3.5.2 Beberapa distribusi peluang peubah acak kontinu 39
- 3.6 Distribusi Sampel Acak 48

4 Penaksiran 54
- 4. 1 Pengantar 54
- 4.2 Penaksir Kemungkinan Maksimum 56
- 4.3 Kriteria Penaksir 61
- 4.4 Penaksiran Parameter 65
- 4.5 Selang Kepercayaan Parameter Rata-rata Populasi 67
  - 4.5.1 Selang kepercayaan $(1 - \alpha)100\%$ rata-rata suatu populasi ($n \geq 30$) 68
  - 4.5.2 Selang kepercayaan $(1 - \alpha)100\%$ rata-rata suatu populasi ($n < 30$) 70
  - 4.5.3 Selang kepercayaan $(1 - \alpha)100\%$ untuk selisih rata-rata dua populasi bebas, $n_1 \geq 30$ dan $n_2 \geq 30$ 72
  - 4.5.4 Selang kepercayaan $(1 - \alpha)100\%$ selisih rataan dua populasi bebas, $n_1 < 30$ dan $n_2 < 30$. Asumsi $\sigma_1^2 = \sigma_2^2$ tapi tidak diketahui 75
  - 4.5.5 Selang kepercayaan $(1 - \alpha)100\%$ selisih rataan dua populasi bebas, $n_1 < 30$ dan $n_2 < 30$. Asumsi $\sigma_1^2 \neq \sigma_2^2$ dan tidak diketahui 78

4.5.6 Selang kepercayaan (1 – α)100% selisih rata-rata pengamatan berpasangan 79

4.6 Selang Kepercayaan Proporsi Populasi 81
4.6.1 Selang kepercayaan (1 – α)100% proporsi sebuah populasi 81
4.6.2 Selang kepercayaan (1 – α)100% selisih proporsi dua populasi 84

4.7 Selang Kepercayaan Variansi 86
4.7.1 Selang kepercayaan (1 – α)100% variansi suatu populasi 86
4.7.2 Selang kepercayaan (1 – α)100% rasio variansi dua populasi 88

5 Pengujian Hipotesis 91
5.1 Konsep Pengujian Hipotesis 91
5.2 Pengujian Hipotesis Rata-rata dan Variansi 102
5.2.1 Uji hipotesis rata-rata suatu populasi ($n \geq 30$) 103
5.2.2 Uji hipotesis rata-rata suatu populasi ($n < 30$) 114
5.2.3 Uji hipotesis selisih rata-rata dua populasi bebas ($n_1 \geq 30$ dan $n_2 \geq 30$) 120
5.2.4 Uji hipotesis selisih rata-rata dua populasi bebas ($n_1 < 30$ dan $n_2 < 30$) 122
5.2.5 Uji hipotesis selisih rata-rata data berpasangan 129
5.2.6 Uji hipotesis variansi suatu populasi 133
5.2.7 Uji hipotesis rasio variansi dua populasi 137
5.3 Uji Hipotesis Proporsi 142
5.3.1 Uji hipotesis proporsi suatu populasi 142
5.3.2 Uji hipotesis selisih proporsi dua populasi 153

6 Aplikasi dalam Penelitian Penulis 157
6.1 Sekilas Analisis Regresi 157
6.2 Artikel Ilmiah Penulis Menggunakan Analisis Regresi 168
6.2.1 Artikel Publikasi *Proceedings of the International Conference on Industrial Engineering and Operations Management* Monterrey, Mexico, November 3-5, 2021 168
6.2.2 Artikel Publikasi *International Journal of Innovation, Creativity and Change*. www.ijicc.net Volume 14, Issue 10, 2020 171
6.2.3 Artikel Publikasi *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management*, Detroit, Michigan, USA, *August* 10 - 14, 2020 175
a *Dark Side of Economic Growth: A Case Study of the Relationship between Economic Growthand Suicide Mortality* 175
b *E-Business Services Strategy with Financial Technology: Evidence from Indonesia* 177
c *Measurement Availability of Clean Water and ElementaryTeachers towards Income of all Districts and Cities in Indonesia* 179

6.3 Uji Kolmogorov-Smirnov untuk Pengujian Kenormalan Data 182

6.4 Artikel Ilmiah Menggunakan Uji Kolmogorov-Smirnov …….. 185
6.4.1 Artikel Publikasi Jurnal Teknik Mesin – ITI Vol.6 No.1, …… 185
Februari 2022
6.4.2 Artikel Publikasi Jurnal IPTEK Volume 1 Nomor 1, April 2015 190

Daftar Pustaka …… 193

DAFTAR TABEL

Tabel 2.1.1	Ruang sampel lantunan 2 dadu sekaligus ……	7
Tabel 2.4.1	Tabel distribusi peluang banyak angka (A) muncul ……	13
Tabel 3.2.1	Tabel distribusi peluang untuk Contoh ……	21
Tabel 3.4.1	Tabel perhitungan ekspektasi ……..	23
Tabel 3.5.1.1	Tabel Binomial untuk n = 15 dan p = 0,6 ……	33
Tabel 3.5.1.2	Tabel hipergeometrik N = 7, n = 3, M = 3 ……..	35
Tabel 3.5.2.1	Tabel Khi-kuadrat dengan $\nu = 9$ untuk beberapa α ……	47
Tabel 4.4.1	Karakteristik Data ……..	65
Tabel 4.5.1.1	Nilai $z_{0,005}$ ……	70
Tabel 4.5.2.1	Statistik data Volume Galon ……..	72
Tabel 4.5.2.2	Nilai $t_{0,025,\nu=7}$ ……	72
Tabel 4.5.3	Nilai $z_{0,02}$ ……	74
Tabel 4.5.4	Nilai $t_{0,02,\ \nu=31}$ ……	77
Tabel 4.5.5	Nilai $t_{0,025\ ,\ \nu\ =\ 21}$ ……..	79
Tabel 4.5.6	Data berpasangan ……	81
Tabel 4.6.2	Nilai $z_{0,05}$ ……	86
Tabel 4.7.1	Statistik panjang badan bayi perempuan ……	87
Tabel 4.7.2.1	Nilai $f_{0,01\ (19,\ 24)}$ ……	89
Tabel 4.7.2.2	Nilai $f_{0,01\ (24,\ 19)}$ ……	89
Tabel 5.2.1	Pengujian hipotesis rata-rata dan variansi ……..	102
Tabel 5.2.1.1	Hasil pengujian hipotesis Contoh 1 ……	108
Tabel 5.2.1.2	Hasil pengujian hipotesis Contoh 2 ……	110
Tabel 5.2.2	Hasil pengujian hipotesis Contoh 4 ……	120
Tabel 5.2.4	Nilai $t_{0,025\ ,\ \nu\ =\ 21}$ ……..	127
Tabel 5.2.5	Data pengamatan berpasangan (Contoh 8) ……	129
Tabel 5.2.6.1	Nilai $\chi^2_{0,025,\nu=9}$ ……..	134
Tabel 5.2.6.2	Nilai $\chi^2_{0,975,\nu=9}$ ……..	134
Tabel 5.2.7.1	Nilai $f_{0,01,\ (19,\ 24)}$ ……	138
Tabel 5.2.7.2	Nilai $f_{0,99,\ (19,\ 24)}$ ……	138
Tabel 5.3.1.1	Nilai peluang Binomial: $P(X \leq 4 \mid p = 0{,}6,\ n = 15)$ ……	143
Tabel 5.3.1.2	Nilai peluang Binomial: $P(X \leq 3 \mid p = 0{,}1,\ n = 15)$ ……	150
Tabel 6.1.1	Pengamatan RLS ……	161
Tabel 6.1.2	Tabel Analisis Variansi (ANOVA) RLS ……	164
Tabel 6.1.3	Uji hipotesis parameter β_i RLS	165
Tabel 6.1.4	Data Model RLB ……	166
Tabel 6.1.5	Uji hipotesis parameter β_i RLB	167
Tabel 6.2.1.1	Nilai Energi Magnetik ……	169
Tabel 6.2.1.2	Hasil uji hipotesis parameter model energi magnetik ……	170
Tabel 6.2.1.3	Prediksi Temperatur 750°C dan waktu 12jam ……	171
Tabel 6.2.1.4	Prediksi Temperatur 850°C dan waktu 12jam ……	171
Tabel 6.2.1.5	Prediksi Temperatur 900°C dan waktu 14jam ……	171
Tabel 6.2.2.1	Statistik deskriptif variabel-variabel *Panel Data Analysis* ……	174
Tabel 6.2.2.2	Hasil uji Hausmann terhadap parameter *Fixed Effect Model* ……	174

Tabel 6.2.3.1	Hasil pengolahan data *Dark Side of Economic Growth* …….	177
Tabel 6.2.3.2	Hasil pengolahan data *Financial Technology* ……	179
Tabel 6.2.3.3	Hasil pengolahan data *Income of all Districts and Cities in Indonesia* …	181
Tabel 6.3	Nilai kritis Uji Kolmogorov-Smirnov ……	184
Tabel 6.4.1.1	Data Lebar Pemotongan (*Kerf*) …….	185
Tabel 6.4.1.2	Statistik Lebar Pemotongan (Kerf)_1 ……	186
Tabel 6.4.1.3	Pengolahan Data Uji Kolmogorov-Smirnov ……..	187
Tabel 6.4.2	Data tingkat intelegensi siswa ……	191

DAFTAR GAMBAR

Gambar 2.1.1	Ruang sampel lantunan 2 koin …….	6
Gambar 2.4.1	Grafik distribusi peluang banyak angka (A) muncul ……..	13
Gambar 3.5.1	Daftar distribusi peluang teoritis pada Program Minitab 19 ……	24
Gambar 3.5.1.1	Langkah 1 menu Distribusi Binomial ……..	26
Gambar 3.5.1.2	Langkah 2 menu Distribusi Binomial ……..	27
Gambar 3.5.1.3	Pilihan menu Distribusi Binomial ……	27
Gambar 3.5.1.4	Input data *probability* ……..	28
Gambar 3.5.1.5	Langkah 3 menu Distribusi Binomial …….	28
Gambar 3.5.1.6	Input data dari kolom pada menu *Probability* …….	29
Gambar 3.5.1.7	Menu *Optional storage* ……..	29
Gambar 3.5.1.8	*Output Probability* Binomial ……..	30
Gambar 3.5.1.9	Langkah 4 menu Distribusi Binomial ……..	30
Gambar 3.5.1.10	Langkah 5 menu Distribusi Binomial ……	31
Gambar 3.5.1.11	*Output Cumulative probability* Binomial ……	31
Gambar 3.5.1.12	Grafik distribusi peluang Binomial, n = 15 dan p = 0,6 …….	32
Gambar 3.5.1.13	Input data Hipergeometrik ……..	35
Gambar 3.5.1.14	Langkah pertama menu Distribusi Hipergeometrik ……	35
Gambar 3.5.1.15	*Output Probability* Hipergeometrik ……	36
Gambar 3.5.1.16	Grafik distribusi peluang Hipergeometrik N = 7, n = 3, M = 3 …..	36
Gambar 3.5.1.17	Kotak dialog *Poisson Distribution* ……..	37
Gambar 3.5.1.18	Input data Poison …….	38
Gambar 3.5.1.19	*Output Probability* Poison ……..	38
Gambar 3.5.1.20	Grafik distribusi peluang Poison $\mu = 4$ …….	38
Gambar 3.5.2.1	Kurva Normal $\mu = 0$, $\sigma^2 = 1$ ……	39
Gambar 3.5.2.2	Kurva Normal $\mu = 1$, $\sigma^2 = 0{,}25$ …….	39
Gambar 3.5.2.3	P($x_1 < X < x_2$) Distribusi Normal …….	40
Gambar 3.5.2.4	Input data Normal ……	41
Gambar 3.5.2.5	Menu *Probability Distributions* dan Normal ……	41
Gambar 3.5.2.6	Input data Normal ……	41
Gambar 3.5.2.7	*Output* Normal $\mu = 800$, $\sigma = 40$ ……..	42
Gambar 3.5.2.8	Ilustrasi z-*score* ……..	42
Gambar 3.5.2.9	*Inverse cumulative probability* $z_{0,025}$ ……..	43
Gambar 3.5.2.10	Grafik Distribusi Khi-kuadrat dengan $\nu = 9$ dan $\nu = 14$ …….	44
Gambar 3.5.2.11	Ilustrasi χ^2_{α} ……..	44
Gambar 3.5.2.12	Input data α Normal ……	45
Gambar 3.5.2.13	Kotak dialog *Calc* …….	45
Gambar 3.5.2.14	*Expression* untuk 1 – α ……..	46
Gambar 3.5.2.15	*Output Expression* 1 – α ……	46
Gambar 3.5.2.16	*Inverse Cumulative Probability* Khi-kuadrat …….	47
Gambar 3.5.2.17	*Output Inverse Cumulative Probability* Khi-kuadrat ……..	47
Gambar 3.5.2.18	Nilai $\chi^2_{0,05,\nu=9}$ ……..…… …	48
Gambar 3.5.2.19	Nilai $\chi^2_{0,995,\nu=9}$ ……..……	48
Gambar 3.6.1	Grafik T dengan df = 10 ……..……..	50

Gambar 3.6.2	Ilustrasi t_{α} …….……..	51
Gambar 3.6.3	Nilai $t_{\alpha,\ \nu = 10}$ untuk beberapa α ……..…….	51
Gambar 3.6.4	Grafik F dengan $\nu_1 = 11$, $\nu_2 = 13$ ……..……	52
Gambar 3.6.5	Ilustrasi f_{α} ……..……..	52
Gambar 3.6.6	Langkah awal menentukan f_{α} dengan Program Minitab …….	53
Gambar 3.6.7	Nilai $f_{\alpha,\ v1 = 11,\ v2 = 13}$ untuk beberapa α …….	53
Gambar 4.5.1	Selang kepercayaan $(1 - \alpha)100\%$ rata-rata suatu populasi ($n \geq 30$) ……	68
Gambar 4.5.2	Selang kepercayaan $(1 - \alpha)100\%$ rata-rata suatu populasi ($n < 30$) …….	71
Gambar 4.6.1.1	Grafik Normal n = 100, p=0,5 ……..…….	82
Gambar 4.6.1.2	Grafik Normal n = 100, p=0,45 ……..…….	82
Gambar 4.6.1.3	Grafik Normal n = 100, p=0,3 ……..…….	82
Gambar 4.6.1.4	Grafik Normal n = 100, p=0,8 ……..…….	82
Gambar 4.7.1	Selang kepercayaan $(1 - \alpha)100\%$ variansi suatu populasi …….	86
Gambar 5.1.1	Daerah penerimaan H_0 dan daerah penolakan H_0 (Binomial) …….	94
Gambar 5.1.2	Daerah penerimaan H_0 dan daerah penolakan H_0 (Normal) ……	96
Gambar 5.1.3	Daerah kritis uji ekaarah $\theta > \theta_0$ Distribusi Normal …….	98
Gambar 5.1.4	Daerah kritis uji ekaarah $\theta < \theta_0$ Distribusi Normal …….	98
Gambar 5.1.5	Daerah kritis uji ekaarah $\theta \neq \theta_0$ Distribusi Normal …….	99
Gambar 5.1.6	Daerah kritis uji ekaarah $\theta > \theta_0$ Distribusi Khi-kuadrat …….	100
Gambar 5.1.7	Daerah kritis uji ekaarah $\theta < \theta_0$ Distribusi Khi-kuadrat …….	100
Gambar 5.1.8	Daerah kritis uji ekaarah $\theta \neq \theta_0$ Distribusi Khi-kuadrat …….	101
Gambar 5.2.1.1	Daerah kritis $Z < -1{,}645$ ……..……..	104
Gambar 5.2.1.2	$z_{sampel} = 17{,}71$ terletak pada daerah kritis ……	104
Gambar 5.2.1.3	Menu 1-*Sample* Z Program Minitab …….	105
Gambar 5.2.1.4	Kotak dialog *One-Sample* Z *for the Mean* …….	106
Gambar 5.2.1.5	Langkah 2 Uji *One-Sample* Z ……..……..	106
Gambar 5.2.1.6	Kotak dialog *One-Sample* Z : *Options*(1) ……..	107
Gambar 5.2.1.7	Kotak dialog *One-Sample* Z : *Options* (2) ……..	107
Gambar 5.2.1.8	*Output* Uji *One-Sample* Z Contoh 1 ……..	108
Gambar 5.2.1.9	Uji hipotesis dengan alternatif *Mean > hypothesized mean* ……..	109
Gambar 5.2.1.10	*Output* Uji *One-Sample* Z Contoh 2 ……	109
Gambar 5.2.1.11	Daerah kritis Contoh 2 ……	110
Gambar 5.2.1.12	Daerah kritis Contoh 3 ……..…….	111
Gambar 5.2.1.13	Uji hipotesis dengan alternatif *Mean ≠ Hypothesized mean* ……	112
Gambar 5.2.1.14	*Output* Uji *One-Sample* Z Contoh 3 ……	112
Gambar 5.2.1.15	z_{sampel} terletak dalam daerah penerimaan H_0 ……	114
Gambar 5.2.2.1	Daerah kritis Contoh 4 …….	115
Gambar 5.2.2.2	Input data pada *worksheet* untuk Contoh 4 ……..	116
Gambar 5.2.2.3	Input data desimal yang salah ……..	116
Gambar 5.2.2.4	Data *Time* atau *Text* tidak muncul pada kotak data …….	117
Gambar 5.2.2.5	Menu 1-*Sample* t Program Minitab …….	117
Gambar 5.2.2.6	Kotak dialog *One-Sample* t *for the Mean* …….	117
Gambar 5.2.2.7	Langkah 2 Uji *One-Sample* t *for the Mean* …….	118
Gambar 5.2.2.8	Hasil Langkah 2 Uji *One-Sample* t *for the Mean* ……	118
Gambar 5.2.2.9	*Perform hyphotesis test* Contoh 4 ……	118

Gambar 5.2.2.10	Kotak dialog *One-Sample* t: *Options*	119
Gambar 5.2.2.11	*Output* Uji *One-Sample* t Contoh 4	119
Gambar 5.2.3.1	Daerah kritis Contoh 5	121
Gambar 5.2.4.1	Daerah kritis Contoh 6	123
Gambar 5.2.4.2	Menu 2-*Sample* t	124
Gambar 5.2.4.3	Kotak dialog *Two-Sample* t *for the Mean*	124
Gambar 5.2.4.4	Kotak dialog *Two-Sample* t: *Options*	125
Gambar 5.2.4.5	Hasil pengujian hipotesis Contoh 6	125
Gambar 5.2.4.6	Daerah kritis Contoh 7	127
Gambar 5.2.4.7	Kotak dialog *Two-Sample* t *for the Mean*	128
Gambar 5.2.4.8	Hasil pengujian hipotesis Contoh 7	128
Gambar 5.2.5.1	Daerah kritis Contoh 8	130
Gambar 5.2.5.2	Input data Contoh 8 pada *worksheet*	131
Gambar 5.2.5.3	Menu *Paired* t (pengamatan berpasangan)	131
Gambar 5.2.5.4	Kotak dialog *Paired* t *for the Mean*	132
Gambar 5.2.5.5	Kotak dialog *Paired* t: *Options*	132
Gambar 5.2.5.6	Hasil pengujian hipotesis Contoh 8	132
Gambar 5.2.6.1	Daerah kritis Contoh 9	134
Gambar 5.2.6.2	Input data Contoh 9	135
Gambar 5.2.6.3	Menu 1 *Variance*	135
Gambar 5.2.6.4	Kotak dialog *One-Sample Variance*	136
Gambar 5.2.6.5	Kotak dialog *Perform hypothesis test One-Sample Variance*	136
Gambar 5.2.6.6	Kotak dialog *One-Sample Variance*: *Options*	136
Gambar 5.2.6.7	Hasil pengujian hipotesis Contoh 9	137
Gambar 5.2.7.1	Daerah kritis Contoh 10	139
Gambar 5.2.7.2	Menu 2 *Variance*	139
Gambar 5.2.7.3	Kotak dialog *Two-Sample Variance*	140
Gambar 5.2.7.4	Input data Contoh 10	140
Gambar 5.2.7.5	Kotak dialog *Two-Sample Variances* : *Options* (1)	140
Gambar 5.2.7.6	Kotak dialog *Two-Sample Variances* : *Options* (2)	141
Gambar 5.2.7.7	Hasil pengujian hipotesis Contoh 10	141
Gambar 5.3.1.1	Menu 1 *Proportion*	144
Gambar 5.3.1.2	Kotak dialog *One-Sample Proportion*	144
Gambar 5.3.1.3	Kotak dialog *One-Sample Proportion* : *Options* (1)	145
Gambar 5.3.1.4	Kotak dialog *One-Sample Proportion* : *Options* (2)	145
Gambar 5.3.1.5	Hasil pengujian hipotesis Contoh 11 (1)	146
Gambar 5.3.1.6	Input data Contoh 1 pada *worksheet*	146
Gambar 5.3.1.7	Opsi *One or more samples, each in a column*	147
Gambar 5.3.1.8	Input data dari kotak data	147
Gambar 5.3.1.9	Data pada kolom C1 Sembuh masuk ke kotak data	148
Gambar 5.3.1.10	Hasil pengujian hipotesis Contoh 11 (2)	148
Gambar 5.3.1.11	Hasil pengujian hipotesis Contoh 11 (3)	149
Gambar 5.3.1.12	Hasil pengujian hipotesis Contoh 12	150
Gambar 5.3.1.13	Input data Contoh 13	151
Gambar 5.3.1.14	Kotak dialog *One-Sample Proportion*: *Options*	152
Gambar 5.3.1.15	Hasil pengolahan data Contoh 13	152

Gambar 5.3.2.1	Menu *Two Proportions* ……..…….	155
Gambar 5.3.2.2	Kotak dialog *Two-Sample Proportions* ……..……	156
Gambar 5.3.2.3	Hasil pengujian hipotesis Contoh 14 ……	156
Gambar 6.3.1	Ilustrasi penentuan nilai-p Uji Kolmogorov-Smirnov	184
Gambar 6.4.1.1	Input data Lebar Pemotongan (Kerf)_1 ……..……	188
Gambar 6.4.1.2	Menu *Normality Test* ……..…….	188
Gambar 6.4.1.3	Kotak dialog *Normality Test* ……..…….	189
Gambar 6.4.1.4	Opsi Uji Kolmogorov-Smirnov ……..…….	189
Gambar 6.4.1.5	Hasil pengujian hipotesis data lebar pemotongan (*kerf*)	190
Gambar 6.4.2	Hasil uji kenormalan data tingkat intelegensi siswa ……	191

1. Pendahuluan

Peranan Statistika dalam pengolahan data pada suatu kegiatan penelitian sangatlah penting. Oleh karenanya terlebih dahulu mari kita pahami salah satu arti Statistika sebagai cabang ilmu Matematika. Statistika adalah ilmu yang mempelajari bagaimana merencanakan, mengumpulkan, menganalisis, menginterpretasi, dan mempresentasikan data. Secara singkat dapat dikatakan bahwa Statistika adalah ilmu yang berkenaan dengan data. Di sini, data dapat diartikan sebagai keterangan atau informasi yang diperlukan untuk mendapatkan gambaran mengenai suatu keadaaan sehingga dimungkinkan untuk memecahkan masalah mengenai keadaan tersebut. Data tersebut dapat berupa angka-angka yang berasal dari perhitungan atau pengukuran kuantitatif atas suatu objek atau penilaian yang bersifat kualitatif atas suatu objek. Berdasarkan kriteria ini, maka ada dua jenis data yaitu :

- Data Kuantitatif : berupa angka hasil pengukuran atau penjumlahan

  Contoh. Skor inteligensi atau IQ (*Intelligence Quotient*) dan Umur Siswa SD.

- Data Kualitatif : bersifat verbal, yang merupakan dekripsi tentang suatu hal, keadaan atau proses

  Contoh. Jenis Kelamin, Kinerja Dosen dan Peringkat Jurnal Nasional.

Angka-angka yang digunakan dalam analisis statistika terdiri atas 4 (empat) kategori, yaitu :

a. Skala nominal, disebut juga skala klasifikasi, merupakan angka yang tidak mempunyai arti hitung. Angka di sini hanya merupakan lambang pengkategorian yang dipergunakan untuk mengklasifikasi suatu benda, sifat, jenis atau orang. Oleh karenanya operasi aljabar matematika tidak berlaku pada angka skala nominal.

   Contoh.

   1. Pengkategorian jenis kelamin yang dapat dibedakan atas pria dan wanita. Data ini dapat diklasifikasikan ke dalam lambang-lambang angka, misalnya pria = 1 dan wanita = 2.
   2. Klasifikasi pekerjaan orangtua mahasiswa, misalnya pegawai negeri = 1, TNI = 2 , pegawai swasta = 3 , wiraswasta = 4.

b. Skala ordinal, disebut juga skala peringkat, adalah angka yang selain berfungsi sebagai lambang pengkategorian, juga mempunyai daya pembeda (pengkategorian di sini telah didasarkan atas hubungan tingkatan), tetapi perbedaan antara angka yang satu dengan angka yang lainnya tidak konstan (tidak mempunyai satuan atau unit perbedaan yang jelas).

Hubungan yang biasa digunakan adalah kategori lebih, seperti lebih besar, lebih tinggi, lebih pintar, dan lain-lain. Sama halnya seperti angka skala nominal, operasi aljabar matematika tidak berlaku pada angka skala ordinal.

Contoh.

Hasil ujian akhir siswa di sebuah SMA menyatakan bahwa : siswa A juara 1, siswa B juara 2 dan siswa C juara 3. Penggunaan angka-angka 1, 2, dan 3 tidak berpedoman bahwa angka yang lebih besar adalah angka yang lebih baik. Dalam contoh ini, tentu saja angka yang lebih kecil menunjukkan peringkat yang lebih baik.

c. Skala interval adalah data yang mempunyai ciri-ciri skala ordinal, tetapi jarak antara tiap angka diketahui dan mempunyai bobot yang sama. Data interval tidak mempunyai angka 0 (nol) mutlak karena angka pada skala ini terletak dalam suatu interval tertentu. Nilai nol pada skala interval tidak berlaku umum untuk suatu pengukuran. Operasi aljabar matematika berlaku pada angka skala interval.

   Contoh.

   1. Nilai ujian mempunyai interval dari 0 sampai dengan 100.

      Di sini mahasiswa yang mempunyai nilai 0 (nol) tidak berarti mahasiswa tersebut tidak mempunyai pengetahuan sama sekali, tetapi mahasiswa tersebut tidak dapat menyelesaikan soal pada saat ujian. Mahasiswa yang memperoleh nilai 80 mempunyai kemampuan yang lebih baik daripada mahasiswa yang memperoleh nilai 30.

   2. Temperatur mempunyai rentangan dari 0°C sampai dengan 100°C.

      Di sini temperatur 0°C bukan berarti udara tidak mempunyai temperatur, sebab 0°C adalah titik beku air. Temperatur udara 15°C sama dengan 0,5 (setengah kali) temperatur udara 30 °C.

d. Skala rasio adalah data yang mempunyai ciri data interval tapi mempunyai angka 0 (nol) mutlak yang berlaku umum untuk suatu pengukuran. Skala rasio merupakan bilangan yang sebenarnya dan tentu saja operasi aljabar matematika juga berlaku pada angka skala interval.

   Contoh

   Berat suatu obyek tertentu, kadar oksigen (saturasi oksigen) , tinggi badan, umur.

Kegiatan pengumpulan data (data mentah) dalam penelitian dapat diperoleh dengan melakukan percobaan (eksperimen), wawancara, melakukan pengamatan (observasi),

menggunakan kuesioner dan lain-lain. Data yang demikian disebut data primer, yaitu data yang dikumpulkan oleh si peneliti sendiri. Namun kita juga dapat memperoleh data melalui sumber-sumber seperti BPS atau dari hasil penelitian orang lain yang kita tidak terlibat di dalamnya. Data yang demikian disebut data sekunder.

Saat ini pengolahan data dengan menggunakan Ilmu Statistika telah diterapkan dalam hampir semua disiplin ilmu, mulai dari ilmu-ilmu alam (seperti Astronomi dan Biologi), ilmu-ilmu Sosial (seperti Ilmu Agama dan Psikologi), ilmu-ilmu di bidang Bisnis dan Ekonomi, hingga Ilmu Kedokteran dan Ilmu Teknik. Statistika juga digunakan dalam penyelenggaraan administrasi pemerintahan. Di Indonesia sendiri ada lembaga pemerintahan yaitu Badan Pusat Statistik (BPS) yang bertugas mengumpulkan dan menyediakan kebutuhan data untuk berbagai macam tujuan. Pada situs web BPS kita bisa memperoleh data-data yang dapat digunakan sebagai data penelitian seperti data Sosial dan Kependudukan, Ekonomi dan Perdagangan, Pertanian dan Pertambangan, dan lain-lain.

Jika didasarkan kepada aktivitas yang dilakukan, maka Statistika dibagi atas:

1. Statistika Deskriptif.

   Aktivitas yang dilakukan dalam Statistika Deskriptif :

   a. Penyajian data mentah ke dalam bentuk yang lebih informatif seperti dalam bentuk tabel distribusi frekuensi, diagram batang (*histogram*) dan kurva (*curve*).
   b. Meringkas kumpulan data ke dalam bentuk pengukuran (bilangan) yang dapat mewakili sekumpulan data tersebut. Dua pengukuran yang penting adalah ukuran pemusatan data dan ukuran penyebaran data. Ukuran pemusatan data adalah nilai atau skor yang dapat digunakan untuk mewakili sekumpulan data. Sedangkan ukuran penyebaran data adalah nilai atau skor yang dapat digunakan untuk menggambarkan bagaimana data menyebar di sekitar ukuran pemusatannya. Ukuran pemusatan data dan ukuran penyebaran data yang paling umum digunakan, masing-masing adalah rata-rata (*mean*) dan standar deviasi (*standard deviation*). Bilangan seperti rata-rata dan standar deviasi ini dikenal sebagai karakteristik suatu kumpulan data.

   Kesimpulan yang diperoleh dari Statistika Deskriptif merupakan kesimpulan yang berlaku hanya terhadap kumpulan data tersebut dan tidak memberikan kesimpulan umum (inferensi) tentang kumpulan yang lebih besar dari data tersebut. Lebih lanjut, data yang kita kumpulkan disebut sampel, sedangkan kumpulan yang lebih besar dari data disebut populasi. Sebagai

contoh, kita melakukan penelitian tentang pengaruh derajat infeksi cacing tanah terhadap tingkat intelegensia siswa Sekolah Dasar (SD) Kelas 4, 5 dan 6 di Indonesia. Maka populasi adalah seluruh siswa SD kelas 4, 5 dan 6 di Indonesia. Namun, jika kita harus mengumpulkan data dari seluruh siswa SD kelas 4, 5 dan 6 di Indonesia, maka dibutuhkan biaya yang tinggi, sumber daya manusia yang besar dan waktu yang lama. Untuk itu kita dapat mengambil sebagian data siswa SD kelas 4, 5 dan 6 di sebuah sekolah. Data ini kita sebut data sampel. Kemudian mengenai karakteristik data. Dalam Statistika, terdapat perbedaan istilah karakteristik untuk populasi dan sampel. Karakteristik populasi disebut parameter sedangkan karakteristik sampel disebut statistik. Semoga dengan pemahaman ini pembaca semakin memahami bahwa Statistika tidak sama dengan statistik. Statistika adalah cabang ilmu Matematika, sedangkan statistik adalah suatu bilangan karakteristik sampel.

2. Statistika Inferensi.

   Di lain pihak, dalam Statistika Inferensi yang dilakukan adalah prosedur pengambilan kesimpulan (inferensi) terhadap data populasi berdasarkan data sampel yang telah dikumpulkan. Dalam Statistika Inferensi dilakukan :

   a. Pendugaan atau penaksiran nilai karakteristik (parameter) populasi.

      Prosedur penaksiran nilai parameter populasi dilakukan dengan menggunakan nilai statistik sampel yang dihitung dengan menggunakan prosedur pada Statistika Deskriptif.

   b. Prosedur pengujian hipotesis terhadap nilai parameter populasi yang diperoleh pada a.
   c. Membuat kesimpulan (inferensi) berdasarkan a dan b.
   d. Pengambilan keputusan atau kebijakan berdasarkan a, b dan c.

   Prosedur a sampai d di atas dilakukan dengan mempertimbangkan resiko terjadinya kesalahan dalam mengambil kesimpulan dan keputusan. Kita dapat menetapkan besarnya resiko kesalahan tersebut dan hal ini dilakukan dengan menggunakan teori peluang. Oleh karenanya pemahaman konsep peluang sangat penting dalam memahami Statistika Inferensi.

Pemaparan tentang Statistika Deskriptif dan Statistika Inferensi merupakan hal yang sangat mendasar dan meski terdapat perbedaan aktivitas yang dilakukan namun keduanya merupakan satu kesatuan dalam aktivitas pengolahan daata, analisa data hingga pengambilan kesimpulan dari suatu kelompok data.

2. Konsep Peluang

Untuk memahami konsep peluang dengan baik, kita akan mulai dengan konsep ruang sampel dan kejadian.

2.1.Ruang Sampel dan Kejadian

Pengertian Ruang Sampel

Ruang sampel adalah himpunan dari semua hasil yang mungkin dalam suatu percobaan, ditulis dengan notasi S dan anggotanya disebut titik sampel. Anggota suatu ruang sampel bisa berhingga atau tak berhingga.

Contoh.

1. Sebuah koin (uang logam) dilantunkan satu kali, hasilnya ada 2 (dua) sisi; misalkan hasilnya sisi angka (sebut A) dan sisi gambar (sebut G), maka : S = { A, G}. Di sini, anggota ruang sampel berhingga, yaitu 2.
2. Sebuah dadu dilantunkan satu kali, hasilnya ada 6 (enam) sisi dadu, maka :

   → S = { •, ••, •••, ••••, •••••, ••••••} = {❶,❷,❸,❹,❺,❻}

   Di sini, anggota ruang sampel berhingga, yaitu 6.
3. Sebuah koin dilantunkan sampai yang muncul adalah Angka, maka :

   → S = { A, GA, GGA, GGGA. ....}

   3 kali lantunan, baru muncul A , ...dan seterusnya

   2 kali lantunan, baru muncul A

   1 kali lantunan, langsung muncul A

   Di sini, anggota ruang sampel bisa tak berhingga, bisa 1 (A saja), bisa 2 (G dan GA), bisa 3 (G, GA, GGA) , atau bisa 100 (artinya, pada lantunan ke-100 baru muncul A) dan seterusnya.

Ruang sampel S adalah suatu himpunan. Sebagaimana halnya suatu himpunan yang mempunyai himpunan bagian, maka ruang sampel pun mempunyai himpunan bagian, yang disebut kejadian (*event*).

Definisi. Kejadian adalah himpunan bagian dari ruang sampel.

Biasanya ditulis dengan notasi A, B, C....

Kejadian juga suatu himpunan, dan karenanya kejadian akan memenuhi operasi-operasi pada himpunan.

Operasi terhadap Kejadian.

Jika S adalah ruang sampel, A dan B himpunan bagian dari S, maka :

A^C = Komplemen kejadian A

$\rightarrow$ Himpunan yang terdiri dari semua anggota S yang bukan anggota A.

$A \cap B$ = Irisan kejadian A dan B

$\rightarrow$ Himpunan yang terdiri dari anggota S yang merupakan anggota A dan juga anggota B.

$A \cup B$ = Gabungan kejadian A dan B

$\rightarrow$ Himpunan yang terdiri dari anggota A saja, atau B saja, atau kedua-duanya.

ϕ = Himpunan kosong, adalah himpunan yang tidak mempunyai anggota.

Contoh.

1. Dalam percobaan lantunan sebuah dadu, jika A adalah kejadian yang berkaitan dengan munculnya mata dadu dengan jumlah bilangan genap dan B adalah kejadian yang berkaitan dengan munculnya mata dadu dengan jumlah kuadrat sempurna bilangan yang lebih kecil atau sama dengan 5, tentukan A dan B.

   Penyelesaian.

   Ruang sampel dari lantunan sebuah dadu adalah

   S = {1, 2, 3, 4, 5, 6}

   Maka : A = {2, 4, 6} dan B = {1, 4}

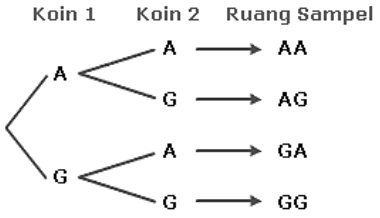
2. Dua buah koin dilantunkan sekaligus, tentukan:

   a. ruang sampel S

   b. kejadian B yang berkaitan dengan yang muncul paling sedikit satu buah gambar

   Penyelesaian

   a. Ruang sampel suatu percobaan dapat juga dinyatakan dalam bentuk diagram pohon atau tabel. Ruang sampel lantunan 2 koin sekaligus dapat dinyatakan dalam bentuk diagram pohon seperti pada Gambar berikut.



Gambar 2.1.1 Ruang sampel lantunan 2 koin

Kejadian yang mungkin :

AA : Muncul sisi angka pada kedua koin

AG : Muncul sisi angka pada koin 1 dan sisi gambar pada koin 2

Maka ruang sampel S = {AA, AG, GA, GG}

Banyak titik sampel ada 4 yaitu AA, AG, GA, dan GG.

b. Kejadian B yang berkaitan dengan yang muncul paling sedikit satu buah gambar :

B = {AA, AG, GA}

3. Dua buah dadu dilantunkan sekaligus, tentukan :

a. ruang sampel S

b. kejadian A yang berkaitan dengan yang muncul jumlahnya adalah 6

c. kejadian B yang berkaitan dengan yang muncul jumlahnya lebih besar dari 11

d. $A \cap B$ e. $A \cup B$

Penyelesaian

a. Ruang sampel percobaan ini ditentukan dengan menggunakan Tabel berikut.

Tabel 2.1.1 Ruang sampel lantunan 2 dadu sekaligus

Dadu 1 / Dadu 2	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)

Sehingga titik sampelnya ada sebanyak 36 kemungkinan.

b. Kejadian A yang berkaitan dengan yang muncul jumlahnya adalah 6 :

A = {(5,1), (4,2), (3,3), (2,4), (1,5)}

c. Kejadian B yang berkaitan dengan yang muncul jumlahnya lebih besar dari 11 :

B = {(6,6)}

d. $A \cap B = \phi$ e. $A \cup B$ = {(5,1), (4,2), (3,3), (2,4), (1,5), (6,6)}

2.2. Peluang suatu Kejadian

Misalkan S adalah suatu ruang sampel yang jumlah anggotanya N, ditulis n(S) = N, dan kejadian A, A $\subset$ S, jumlah anggota A, n(A) = n. Peluang terjadinya kejadian A, ditulis P(A), mengikuti hukum-hukum berikut :

(i) $0 \leq P(A) \leq 1$ (ii) $P(A) = n / N$ (iii) $P(S) = 1$

(iv) $P(A^C) = 1 - P(A)$ (v) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

<u>Catatan.</u> Sampel Acak

Dalam Statistika dikenal istilah sampel acak. Ini merupakan salah satu prosedur paling sederhana dalam penarikan sampel dari suatu populasi. Dalam konteks kejadian dan ruang sampel, sampel acak adalah cara mengambil atau memilih sebuah titik sampel dalam ruang sampel, di mana setiap anggota ruang sampel mempunyai kemungkinan atau peluang yang sama untuk terpilih. Berikut ini percobaan-percobaan yang anggota-anggota ruang sampelnya mempunyai kemungkinan atau peluang yang sama untuk terjadi.

1. Dari lantunan sebuah koin yang seimbang.
   Pemahaman seimbang bermakna bahwa setiap sisi koin mempunyai peluang yang sama untuk muncul ketika koin dilantunkan.
   Di sini ada 2 anggota S = {G, A} $\rightarrow$ P(G) = P(A) = ½
2. Dari lantunan sebuah dadu yang seimbang.
   Pemahaman seimbang bermakna bahwa setiap sisi dadu mempunyai peluang yang sama untuk muncul ketika koin dilantunkan.
   Ada 6 anggota S, maka peluang masing-masing sisi dadu untuk muncul adalah 1/6.
3. Jika sebuah koin yang seimbang dilantunkan 2 kali, S = {GG,GA,AG,AA}, maka peluang masing-masing anggota untuk terjadi adalah ¼.
4. Jika sebuah koin yang seimbang dilantunkan 3 kali dengan S = {GGG, GGA, GAG, AGG, GAA, AGA, AAG, AAA}, jadi n(S) = 8, maka peluang masing-masing anggota untuk terjadi adalah 1/8.
5. Dalam sebuah kotak terdapat 3 bola merah (sebut M1, M2 dan M3) yang bentuk dan ukurannya sama sehingga tidak dapat dibedakan kecuali jika dilihat. Jika kita mengambil sebuah bola secara acak tanpa melhat ke dalam kotak, maka peluang masing-masing bola untuk terambil adalah 1/3.

Contoh.

1. Kepala pabrik mengatakan bahwa dari 100 barang yang diproduksinya, ada 25 yang rusak. Jika barang dibungkus rapi, kemudian seorang pembeli mengambil satu barang secara acak, berapakah peluang bahwa barang tersebut rusak? Berapakah peluang bahwa barang tersebut tidak rusak?

   Penyelesaian

   Dalam ruang sampel ini n(S) = 100, dan ada 2 kemungkinan yang terjadi, yaitu :

   - barang yang dibeli rusak
   - barang yang dibeli tidak rusak

   Misalkan A adalah kejadian bahwa barang yang dibeli rusak ; dengan n(A) = 25

   $$\rightarrow P(A) = \frac{n(A)}{n(S)} = \frac{25}{100} = \frac{1}{4}$$

   Dan kejadian bahwa barang yang dibeli tidak rusak dinyatakan sebagai A^c ; $n(A^c) = 75$

   $$\rightarrow P(A^c) = 1 - P(A) = 1 - \frac{1}{4} = \frac{3}{4} \quad \text{atau} \quad P(A^c) = \frac{n(A^c)}{n(S)} = \frac{75}{100} = \frac{3}{4}$$

2. Dalam sebuah kotak terdapat 4 bola merah dan 6 bola putih yang bentuk dan ukurannya sama. Jika anda mengambil sebuah bola tanpa melihat ke dalam kotak, berapakah peluang bahwa bola tersebut berwarna merah? Berapakah peluang bahwa bola tersebut berwarna putih?

   Penyelesaian

   Dalam ruang sampel ini n(S) = 10, dan ada 2 kemungkinan yang terjadi, yaitu :

   - bola yang terambil berwarna merah
   - bola yang terambil berwarna putih

   Misalkan A adalah kejadian bahwa bola yang terambil berwarna merah ; dengan n(A) = 4

   $$\rightarrow P(A) = \frac{n(A)}{n(S)} = \frac{4}{10} = 0{,}4$$

   Dan B kejadian bahwa bola yang terambil berwarna putih ; n(B) = 6

   $$\rightarrow P(B) = \frac{n(B)}{n(S)} = \frac{6}{10} = 0{,}6$$

2.3. Perhitungan Jumlah Titik Sampel

Dalam penentuan peluang suatu kejadian, sangatlah penting untuk mengetahui banyaknya titik sampel dalam suatu percobaan. Terutama jika suatu percobaan dilakukan dalam beberapa langkah dan setiap langkah menghasilkan berbagai kemungkinan yang berbeda, maka diperlukan suatu cara untuk menghitung seluruh hasil yang mungkin dari rangkaian langkah tersebut (banyaknya titik sampel/anggota ruang sampel atau n(S)). Berikut ini diberikan dua cara untuk menentukan banyaknya titik sampel.

a. Aturan Perkalian. Jika suatu percobaan dilakukan dalam k langkah, langkah pertama menghasilkan n_1 kemungkinan, langkah kedua menghasikan n_2 kemungkinan, dan seterusnya, langkah ke-k menghasilkan n_k kemungkinan, maka percobaan menghasilkan :

   $n_1 . n_2 . .... n_k$ kemungkinan.

   Dengan perkataan lain, dalam percobaan ini : n(S) = **$n_1 . n_2 . .... n_k$**

Contoh.

1. Percobaan lantunan sebuah koin sebanyak 3 kali, dengan masing-masing lantunan mempunyai 2 kemungkinan, maka : $n(S) = 2.2.2 = 2^3$.
2. Percobaan lantunan sebuah dadu sebanyak 2 kali, dengan masing-masing lantunan mempunyai 6 kemungkinan, maka : $n(S) = 6.6 = 6^2$

b. Kombinasi.

   Kadang kala di dalam prakteknya, pemilihan sampel dilakukan dengan menarik sejumlah objek dari sekumpulan objek yang ada. Misalnya kita mengambil 2 bola secara acak dari sebuah kotak yang berisi 6 bola (sebut M1 hingga M6) yang bentuk dan ukurannya sama. Maka untuk menentukan berapa banyak kemungkinan terjadinya kombinasi 2 bola, apakah M1 dan M2, apakah M2 dan M5 dan seterusnya; digunakan rumus kombinasi berikut.

   Banyaknya kombinasi n objek jika diambil sebanyak x :

$$ {}_nC_x = \binom{n}{x} = \frac{n!}{x!(n-x)!} $$

   di mana:

   $n ! = n . (n – 1) . (n – 2) . ... . 3 . 2 . 1$

Contoh.

Dalam sebuah kotak terdapat 4 bola berwarna merah yang bentuk dan ukurannya sama sehingga tidak dapat dibedakan jika tidak dilihat. Sebut bola M1, M2, M3 dan M4.

(i) Jika 1 bola diambil secara acak tanpa melihat ke dalam kotak, berapa banyak kemungkinan yang terjadi?

(ii) Jika 2 bola diambil secara acak tanpa melihat ke dalam kotak dan tanpa pengembalian, berapa banyak kemungkinan yang terjadi?

Penyelesaian.

(i) Di sini n = 4 dan x = 1, maka banyak kemungkinan yang terjadi:

$${}_4C_1 = \binom{4}{1} = \frac{4!}{1!(4-1)!} = 4$$ yaitu bola M1, M2, M3 dan M4. Jadi n(S) = 4.

(ii) Di sini n = 4 dan x = 2, maka banyak kemungkinan yang terjadi:

$${}_4C_2 = \binom{4}{2} = \frac{4!}{2!(4-2)!} = 6$$

yaitu bola M1 dan M2, M1 dan M3, M1 dan M4, M2 dan M3, M2 dan M4, M3 dan M4.

Jadi n(S) = 6.

c. Kombinasi dan Aturan Perkalian.

Kadang kala penentuan banyaknya n(S) menggunakan rumus kombinasi dan aturan perkalian, seperti yang diberikan pada contoh berikut.

Contoh.

Sebuah kotak berisi 6 bola merah (M1, M2, M3, M4, M5 dan M6) dan 4 bola putih (P1, P2, P3 dan P4) yang semuanya bentuk dan ukurannya sama sehingga tidak bisa dibedakan jika tidak dilihat. Apabila 3 bola diambil secara acak dan tanpa pengembalian, hitung banyaknya titik sampel jika yang terambil :

(i) 2 merah dan 1 putih

(ii) Ketiga-tiganya merah

Penyelesaian.

Dalam soal ini percobaan dilakukan dalam 2 langkah karena warna bola ada 2. Oleh karenanya akan digunakan rumus kombinasi dan aturan perkalian.

(i) Banyaknya kemungkinan terambilnya bola merah sebanyak 2 bola : ${}_6C_2 = \binom{6}{2}$

Banyaknya kemungkinan terambilnya bola putih sebanyak 1 bola : ${}_4C_1 = \binom{4}{1}$

Maka banyaknya kemungkinan terambilnya 2 bola merah dan 1 bola putih adalah

$\binom{6}{2} \cdot \binom{4}{1} \rightarrow$ menggunakan Aturan Perkalian

(ii) Dalam kasus kedua semua bola berwarna merah. Berarti tidak ada bola putih yang terambil.

Maka banyaknya kemungkinan terambilnya 3 bola merah dan 0 bola putih adalah

$\binom{6}{3} \cdot \binom{4}{0} \rightarrow$ menggunakan Aturan Perkalian

2.4. Distribusi Peluang dan Grafik

Dari suatu percobaan kita dapat menentukan distribusi peluang dari suatu kejadian yang kita amati. Untuk menjelaskannya, tinjau lantunan sebuah koin sebanyak 3 kali, maka ruang sampel dari percobaan ini adalah

S = {GGG, GGA, GAG, AGG, GAA, AGA, AAG, AAA}.

Misalkan kita mengamati kejadian X yaitu banyaknya A yang muncul, maka nilai-nilai :

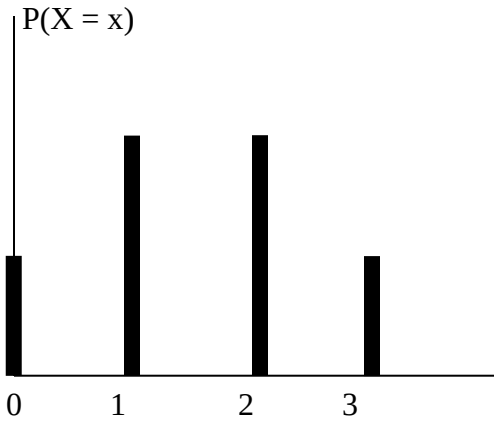
X = 0 $\rightarrow$ berkaitan dengan 1 kejadian munculnya GGG ; P(X=0) = 1/8

X = 1 $\rightarrow$ berkaitan dengan 3 kejadian munculnya GGA, GAG, AGG ; P(X=1) = 3/8

X = 2 $\rightarrow$ berkaitan dengan 3 kejadian munculnya GAA, AGA, AAG ; P(X=2) = 3/8

X = 3 $\rightarrow$ berkaitan dengan 1 kejadian munculnya AAA ; P(X=3) = 1/8

Nilai-nilai P(X = 0), P(X = 1), P(X = 2) dan P(X = 3) dapat kita ringkas dalam bentuk grafik (Gambar 2.4.1) distribusi peluang atau tabel distribusi peluang (Tabel 2.4.1), seperti berikut ini.



Gambar 2.4.1 Grafik distribusi peluang banyak angka (A) muncul

X	P(X = x)
0	1/8
1	3/8
2	3/8
3	1/8
Total	1

Tabel 2.4.1 Tabel distribusi peluang banyak angka (A) muncul

2.5. Peluang Bersyarat, Aturan Perkalian dan Aturan Penjumlahan

Sekarang kita masuk ke dalam konsep di mana perhatian kita tidak saja terhadap satu kejadian tetapi terhadap dua atau lebih kejadian yang mungkin terjadi secara bersamaan. Dalam topik ini kita akan menuju kepada istilah :

a. Saling meniadakan. Pengertian dua kejadian saling meniadakan adalah dua kejadian yang tidak mungkin terjadi secara bersamaan. Jika kejadian yang satu sudah terjadi, maka kejadian yang lain tidak mungkin terjadi.
b. Saling bebas. Pengertian dua kejadian saling bebas adalah dua kejadian dapat terjadi secara bersamaan tetapi terjadinya kejadian yang satu tidak mempengaruhi terjadinya kejadian yang lain.

Pemahaman konsep peluangnya terdiri dari peluang bersyarat, aturan perkalian dan aturan penjumlahan. Kita mulai dengan definisi peluang bersyarat.

Definisi. (Menghitung Peluang Bersyarat) Peluang terjadinya suatu kejadian A dengan diketahui bahwa kejadian B sudah terjadi disebut peluang bersyarat A|B, ditulis dan didefinisikan sebagai :

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

di mana

$P(A \cap B)$ peluang terjadinya kejadian A dan B

$P(A)$ peluang terjadinya kejadian A

$P(B)$ peluang terjadinya kejadian B

Dari rumus ini diperoleh :

$P(A \cap B) = P(B)\, P(A|B)$ yang merupakan peluang terjadinya kejadian A dan kejadian B

Contoh.

1. Dalam sebuah kotak terdapat 6 bola merah yang bentuk dan ukurannya sama sehingga tidak bisa dibedakan jika tidak dilihat. Jika 2 bola diambil secara acak tanpa pengembalian, berapakah peluang terambilnya kedua bola merah tersebut?

   Penyelesaian.

   Misalkan M_1 adalah kejadian terambilnya bola merah pada pengambilan pertama, peluangnya, ditulis $P(M_1)$ dan nilainya adalah

   $$P(M_1) = \frac{1}{6}$$

   dan M_2 adalah kejadian terambilnya bola merah pada pengambilan kedua (tanpa mengembalikan bola merah yang diambil pada pengambilan pertama, yang berarti sekarang bola di dalam kotak tinggal 5), maka peluang terjadinya ditulis sebagai $P(M_2 \mid M_1)$, dan nilainya adalah

   $$P(M_2 \mid M_1) = \frac{1}{5}$$

   Maka peluang terambilnya kedua bola merah adalah

   $$P(M_1 \cap M_2) = P(M_1)\, P(M_2 \mid M_1) = \frac{1}{6} \cdot \frac{1}{5} = \frac{1}{30}$$

2. Dalam sebuah kotak terdapat 4 bola merah dan 6 bola putih yang bentuk dan ukurannya sama. Jika anda mengambil dua buah bola tanpa melihat ke dalam kotak dan tanpa pengembalian, berapakah
   a. peluang bahwa kedua bola tersebut berwarna merah?
   b. peluang bahwa kedua warna terambil ?

Penyelesaian

Ruang sampel percobaan ini adalah 4 bola merah dan 6 bola putih, n(S) = 10.

a. Misalkan M_1 adalah kejadian terambilnya bola merah pada pengambilan pertama, dan peluangnya, ditulis $P(M_1)$ adalah

   $$P(M_1) = \frac{n(M_1)}{n(S)} = \frac{4}{10}$$

Misalkan M_2 adalah kejadian terambilnya bola merah pada pengambilan kedua dan karena bola merah yang terambil pada pengambilan pertama tidak dikembalikan, maka dalam kotak sekarang tinggal 9 bola, 3 merah dan 6 putih – sehingga ruang sampel sekarang mempunyai anggota tinggal 9, sebut ruang sampel yang baru ini S' , maka $n(S') = 9$. Juga karena M_2 terjadi setelah M_1 terjadi, maka peluang terjadinya ditulis sebagai $P(M_2 \mid M_1)$, dan nilainya adalah

$$P(M_2 \mid M_1) = \frac{n(M_2 \mid M_1)}{n(S')} = \frac{3}{9}$$

Maka peluang bahwa kedua bola tersebut berwarna merah ditulis $P(M_1 \cap M_2)$, yaitu

$$P(M_1 \cap M_2) = P(M_1)\, P(M_2 \mid M_1) = \frac{4}{10} \cdot \frac{3}{9} = \frac{2}{5} \cdot \frac{1}{3} = \frac{2}{15}$$

b. Kejadian kedua warna terambil dapat terjadi dalam 2 kemungkinan, yaitu :

- Merah terambil pada pengambilan pertama dan putih terambil pada pengambilan kedua yaitu M_1 dan P_2 atau $M_1 \cap P_2$
- Putih terambil pada pengambilan pertama dan merah terambil pada pengambilan kedua yaitu P_1 dan M_2 atau $P_1 \cap M_2$

Jadi peristiwa kedua warna terambil ditulis :

(M_1 dan P_2) atau (P_1 dan M_2) ditulis : $(M_1 \cap P_2) \cup (P_1 \cap M_2)$

Dan untuk menghitung peluangnya harus digunakan hukum peluang dan pengertiannya sebagai berikut :

Hukum peluang ke (v) : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Untuk menggunakan hukum ini, misalkan :

Kejadian A = $M_1 \cap P_2$ dan Kejadian B = $P_1 \cap M_2$

Maka :

$$P[(M_1 \cap P_2) \cup (P_1 \cap M_2)] = P(M_1 \cap P_2) + P(P_1 \cap M_2) - P[(M_1 \cap P_2) \cap (P_1 \cap M_2)]$$

$P(M_1 \cap P_2)$ ditentukan seperti pada jawaban bagian a di atas, yaitu :

$$P(M_1 \cap P_2) = P(M_1)\, P(P_2 \mid M_1) = \frac{4}{10} \cdot \frac{6}{9} = \frac{2}{5} \cdot \frac{2}{3} = \frac{4}{15}$$

Dan : $P(P_1 \cap M_2) = P(P_1)\, P(M_2 \mid P_1) = \frac{6}{10} \cdot \frac{4}{9} = \frac{3}{5} \cdot \frac{4}{9} = \frac{12}{45} = \frac{4}{15}$

Sedangkan : $P[(M_1 \cap P_2) \cap (P_1 \cap M_2)] = 0$, karena kejadian $(M_1 \cap P_2)$ dan $(P_1 \cap M_2)$ tidak mungkin terjadi secara bersamaan, karena kita hanya mengambil 2 buah bola, sedangkan kejadian $(M_1 \cap P_2)$ dan $(P_1 \cap M_2)$ berarti kita mengambil 4 buah bola. Jadi, jika 2 buah kejadian tidak mungkin terjadi secara bersamaan, maka peluang terjadinya adalah nol.

Dengan demikian :

$$P[(M_1 \cap P_2) \cup (P_1 \cap M_2)] = \frac{4}{15} + \frac{4}{15} - 0 = \frac{8}{15}$$

Rumus : $P(A \cap B) = P(B)\, P(A|B)$ dapat diperluas untuk 3 kejadian A, B dan C. Misalkan kejadian yang pertama terjadi adalah kejadian A ; kemudian kejadian B terjadi setelah kejadian A terjadi, dan terakhir kejadian C terjadi setelah kejadian A dan B terjadi (A dan B ditulis : $A \cap B$), maka :

$$P(A \cap B \cap C) = P(A) \,.\, P(B|A) \,.\, P(C|\, A \cap B)$$

Contoh. Dalam sebuah kotak terdapat 4 bola berwarna merah, 5 bola berwarna kuning dan 6 bola berwarna biru. Jika kita hendak mengambil 3 bola secara acak tanpa pengembalian, hitung peluang bahwa yang terambil pertama kali adalah bola merah, kedua warna kuning dan ketiga warna biru.

Penyelesaian

Misalkan :

M_1 adalah kejadian terambilnya bola merah pada pengambilan pertama

K_2 adalah kejadian terambilnya bola merah pada pengambilan kedua

B_3 adalah kejadian terambilnya bola merah pada pengambilan ketiga

Maka :

$$P(M_1 \cap K_2 \cap B_3) = P(M_1) \,.\, P(K_2|M_1) \,.\, P(B_3|\, M_1 \cap K_2) = \frac{4}{15}.\frac{5}{14}.\frac{6}{13}$$

Kejadian yang Saling Meniadakan (*Mutually Exclusive*)

Jika kejadian A tidak bisa terjadi secara bersamaan dengan kejadian B, artinya jika A sudah terjadi B tidak akan terjadi atau sebaliknya, maka A dan B dikatakan saling meniadakan. Dengan demikian kejadian $A \cap B$ tidak mungkin terjadi, sehingga : $P(A \cap B) = 0$.

Kejadian yang saling meniadakan membuat hukum ke (v) peluang menjadi :

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

Contoh.

1. Jika Anda mengambil satu kartu secara acak dari satu set kartu bridge. A kejadian dapat kartu As dan B kejadian dapat kartu wajit . Hitung P(A $\cup$ B) .

   Penyelesaian. Perhatikan, bahwa dalam soal ini kita hendak mengambil hanya 1 kartu! Jadi kejadian terambilnya kartu As dan B kejadian dapat kartu wajit (As $\cap$ Wajit) atau (A $\cap$ B) mungkin terjadi, dengan perkataan lain kita dapat menghitung :

   $$P(As \cap Wajit) = \frac{n(As \cap Wajit)}{n(kartu)} = \frac{1}{52} = P(A \cap B)$$

   Dan :

   P(A) = peluang terambilnya kartu As = $\frac{n(As)}{n(kartu)} = \frac{4}{52}$

   P(B) = peluang terambilnya kartu wajit = $\frac{n(wajit)}{n(kartu)} = \frac{13}{52}$

   Maka : $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$

   Jadi kedua kejadian ini bukan kejadian yang saling meniadakan; karena mungkin saja kita mendapatkan sebuah kartu berwarna merah dan kartu itu adalah kartu wajit !

2. Jika Anda mengambil satu kartu secara acak dari satu set kartu bridge. A kejadian terambil kartu hati dan B kejadian terambil kartu wajit. Hitung P(A $\cup$ B).

   Penyelesaian. Perhatikan, bahwa dalam soal ini kita hendak mengambil hanya 1 kartu! Jadi kejadian terambilnya kartu hati dan B kejadian terambilnya kartu wajit (Hati $\cap$ Wajit) tidak mungkin terjadi secara bersamaan, dengan perkataan lain : P(Hati $\cap$ Wajit) = 0, sehingga :

   $$P(Hati \cup Wajit) = P(Hati) + P(Wajit)$$

   Di mana :

   P(Hati) = peluang terambilnya kartu hati = $\frac{n(kartu\ hati)}{n(kartu)} = \frac{13}{52} = \frac{1}{4}$

   P(Wajit) = peluang terambilnya kartu wajit = $\frac{n(kartu\ wajit)}{n(kartu)} = \frac{13}{52} = \frac{1}{4}$

   $P(Hati \cup Wajit) = P(Hati) + P(Wajit) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

Perhatikan, kedua kejadian ini adalah dua kejadian yang saling meniadakan, karena tidak mungkin kita mengambil dua kartu sekaligus yaitu kartu hati dan kartu wajit; karena kita hanya ingin mengambil 1 kartu!

Kejadian yang Saling Bebas (*Independent Events*)

Dalam topik Peluang Bersyarat, kita mempunyai rumus :

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

yaitu peluang terjadinya 2 kejadian A dan B, di mana kejadian B terjadi terlebih dahulu, barulah kejadian A terjadi. Rumus ini digunakan apabila terjadinya kejadian A dipengaruhi oleh telah terjadinya kejadian B, disebut juga bahwa A dan B tidak bebas.

Dua kejadian atau lebih dikatakan kejadian yang saling bebas apabila terjadinya kedua kejadian tersebut tidak saling mempengaruhi. Akibatnya : $P(A|B) = P(A)$.

Definisi. Jika A dan B adalah 2 kejadian yang saling bebas maka :

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

Catatan

- Kejadian yang tidak bebas umumnya terjadi jika dalam pengambilan sampel acak dilakukan tanpa pengembalian (*without replacement*).
- Kejadian yang saling bebas umumnya terjadi jika dalam pengambilan sampel acak dilakukan dengan pengembalian (*with replacement*).

Rumus peluang untuk 2 kejadian yang saling bebas : $P(A \cap B) = P(A) \,.\, P(B)$; dapat diperluas untuk n kejadian yang saling bebas, yaitu jika kejadian-kejadian A_1 , A_2 , ... , A_n adalah kejadian-kejadian yang saling bebas, maka :

$$P(A_1 \cap A_2 \cap .... \cap A_n) = P(A_1) \,.\, P(A_2) ..... P(A_n)$$

Contoh.

1. Jika Anda mengambil secara acak 2 kartu berturut-turut dari suatu set kartu bridge. Berapa peluang bahwa pengambilan pertama dan kedua adalah kartu As, pengambilan tanpa pengembalian.

   Penyelesaian.

   Perhatikan, kita hendak mengambil 2 kartu berturut-turut dan tanpa pengembalian, maka :

$$P(As_1 \cap As_2) = P(As_1)\, P(As_2|\, As_1) = \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}$$

2. Jika Anda mengambil secara acak 3 kartu berturut-turut dari suatu set kartu bridge. Berapa peluang bahwa pengambilan pertama, kedua dan ketiga adalah kartu wajit (*diamond*), pengambilan tanpa pengembalian.

   Penyelesaian.

   Perhatikan, kita hendak mengambil 3 kartu berturut-turut dan tanpa pengembalian, maka :

$$P(W_1 \cap W_2 \cap W_3) = P(W_1)\, P(W_2|\, W_1)\, P(W_3|\, W_1 \cap W_2) = \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50}$$

3. Satu bola diambil secara acak dari satu kotak yang berisi 6 bola merah, 4 putih dan 5 biru. Berapa peluang bahwa bola yang terambil pertama adalah merah, kedua putih dan ketiga biru, bila :

   a. pengambilan tanpa pengembalian.

   b. pengambilan dengan pengembalian.

   Penyelesaian.

   a. Pengambilan tanpa pengembalian.

$$P(M_1 \cap P_2 \cap B_3) = P(M_1)\, P(P_2|\, M_1)\, P(B_3|\, M_1 \cap P_2) = \frac{6}{15} \cdot \frac{4}{14} \cdot \frac{5}{13}$$

   b. Pengambilan dengan pengembalian.

$$P(M_1 \cap P_2 \cap B_3) = P(M_1)\, P(P_2|\, M_1)\, P(B_3|\, M_1 \cap P_2)$$

$$= P(M_1)\, P(P_2)\, P(B_3) = P(M)\, P(P) P(B) = \frac{6}{15} \cdot \frac{4}{15} \cdot \frac{5}{15}$$

3. Distribusi Peluang Teoritis

Dalam Pasal 2 kita telah membahas sekilas konsep dasar peluang. Kita mencoba memahami bagaimana menghitung peluang melalui percobaan-percobaan sederhana yang sering ditemukan dalam kehidupan sehari-hari. Kelak, akan kita pahami juga bahwa percobaan-percobaan sederhana tersebut dapat digeneralisasi dalam bentuk rumus-rumus peluang yang disebut rumus distribusi peluang teoritis. Untuk memahami hal itu terlebih dahulu kita harus memahami konsep peubah acak atau variabel random.

3.1. Peubah Acak

Definisi. Peubah acak (*Random variables*) adalah suatu fungsi yang didefinisikan pada ruang sampel S, dan biasanya ditulis dengan huruf seperti X, Y dan Z; sedangkan nilai peubah acak terkait ditulis dengan huruf x, y dan z. Peubah Acak terdiri atas :

- Peubah acak kontinu, berasal dari nilai yang diukur dan terletak dalam suatu selang/interval.
- Peubah acak diskrit, berasal dari nilai yang dihitung/dicacah dan nilainya 0, 1, 2, ...

Contoh.

1. Dari Lantunan sebuah koin sebanyak 3 kali, S = {GGG, GGA, GAG, AGG, GAA, AGA, AAG, AAA}. Misalkan X = peubah acak yang menyatakan banyaknya A yang muncul, maka nilai-nilai :

   X = 0 $\rightarrow$ berkaitan dengan kejadian munculnya GGG ; P(X=0) = 1/8

   X = 1 $\rightarrow$ berkaitan dengan kejadian munculnya GGA, GAG, AGG ; P(X=1) = 3/8

   X = 2 $\rightarrow$ berkaitan dengan kejadian munculnya GAA, AGA, AAG ; P(X=2) = 3/8

   X = 3 $\rightarrow$ berkaitan dengan kejadian munculnya AAA ; P(X=3) = 1/8

   $\rightarrow$ PA X di atas merupakan PA Diskrit, nilai-nilai X diperoleh dari mencacah.

2. Peubah acak X yang menyatakan Tinggi Badan, Berat Badan atau Nilai Ujian Tengah Semester merupakan contoh-contoh peubah acak kontinu, karena nilai-nilai peubah acak ini terletak dalam suatu interval/selang.

Penentuan distribusi peluang suatu peubah acak ditentukan oleh apakah peubah acak tersebut peubah acak diskrit atau peubah acak kontinu. Kita mulai dengan konsep distribusi peluang peubah acak diskrit.

3.2. Distribusi Peluang Peubah Acak Diskrit

Definisi. Misalkan X adalah suatu peubah acak diskrit. Fungsi peluang peubah acak diskrit disebut fungsi massa peluang (*probability mass function* = pmf) yang memenuhi definisi berikut:

1. $P(X= x) = f(x)$ 2. $0 \le f(x) \le 1$ 3. $\sum_x f(x) = 1$

4. $P(a \le X \le b) = \sum_{x=a}^{b} f(x) = f(a) + f(a+1) + ... + f(b-1) + f(b)$

5. Distribusi kumulatif X, diberi notasi F(x), didefinisikan oleh :

$$F(x) = P(X \le x) = \sum_{t \le x} f(t)$$

Distribusi peluang dapat disajikan dalam bentuk grafik yang disebut diagram batang, yang merupakan pasangan (x, f(x)), sedangkan grafik distribusi kumulatif merupakan fungsi tangga.

Contoh.

Dari percobaan lantunan sebuah koin sebanyak 3 kali. Jika X banyaknya angka yang muncul.

a. Buatlah tabel distribusi peluang f(x)
b. Hitung $P(X \le 2)$
c. Hitung $P(1 < X \le 2)$
d. Hitung $P(X \ge 1)$

Penyelesaian

a. Dari Contoh sebelumnya, kita peroleh :

$X = 0 \rightarrow$ berkaitan dengan kejadian munculnya GGG ; $P(X=0) = 1/8$

$X = 1 \rightarrow$ berkaitan dengan kejadian munculnya GGA, GAG, AGG ; $P(X=1) = 3/8$

$X = 2 \rightarrow$ berkaitan dengan kejadian munculnya GAA, AGA, AAG ; $P(X=2) = 3/8$

$X = 3 \rightarrow$ berkaitan dengan kejadian munculnya AAA ; $P(X=3) = 1/8$

Nilai-nilai di atas dapat kita sajikan dalam Tabel distribusi peluang sebagai berikut :

Tabel 3.2.1 Tabel distribusi peluang untuk Contoh

x	0	1	2	3
P(X = x) = f(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

b. $P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{8} + \frac{3}{8} + \frac{1}{8} = \frac{7}{8}$

c. $P(1 < X \le 2) = P(X = 2) = \frac{3}{8}$ d. $P(X \ge 1) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{7}{8}$

3.3. Distribusi Peluang Peubah Acak Kontinu

Sebelum kita masuk ke konsep distribusi peluang peubah acak kontinu, perlu dipahami bahwa seperti yang dijelaskan pada awal pasal ini bahwa peubah acak kontinu berasal dari nilai yang diukur dan terletak dalam suatu selang/interval. Hal ini bermakna bahwa nilai-nilai peubah acak kontinu terletak dalam suatu selang. Konsekuensinya peluang terjadinya hanya pada suatu selang dan peluang untuk x tertentu adalah nol, yang ditulis P(X = x) = f(x) = 0.

Definisi. Fungsi peluang peubah acak kontinu disebut fungsi padat peluang (*probability density function* = pdf) yang memenuhi definisi berikut :

1. $P(X = x) = f(x) = 0$
2. $0 \le f(x) \le 1 \; ; \; \forall x \in \Re$
3. $\int_{-\infty}^{\infty} f(x)\,dx = 1$
4. $P(a < X < b) = \int_{a}^{b} f(x)\,dx$
5. Distribusi kumulatif :

   $F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt$
6. $P(a < X < b) = F(b) - F(a)$

Contoh.

PA kontinu X mempunyai fungsi padat peluang : $f(x) = \begin{cases} x^2/3 & , \; -1 < x < 2 \\ 0 & , \text{ untuk } x \text{ lain} \end{cases}$

a. Hitung $P(0 < X < 1)$.

b. Tentukan F(x) dan hitung $P(0 < X < 1)$ dengan menggunakan F(x).

Penyelesaian

a. $P(0 < X < 1) = \int_{0}^{1} \frac{x^2}{3}\,dx = \frac{1}{9}x^3 \Big|_{0}^{1} = \frac{1}{9}$

b. $F(x) = \int_{-\infty}^{x} \frac{t^2}{3}\,dt = \int_{-1}^{x} \frac{t^2}{3}\,dt = \frac{1}{9}t^3 \Big|_{-1}^{x} = \frac{1}{9}[x^3 - (-1)^3] = \frac{1}{9}(x^3 + 1)$

   Maka : $P(0 < X < 1) = F(1) - F(0) = \frac{1}{9}\left\{(1^3 + 1) - (0^3 + 1)\right\} = \frac{1}{9}$

3.4. Ekspektasi (Nilai Harapan atau Harapan Matematik)

Definisi. Misalkan X suatu peubah acak dengan distribusi peluang f(x). Nilai harapan X atau harapan matematik X atau rata-rata X, ditulis $\mu_X = E[X]$, didefinisikan sebagai:

$E[X] = \sum_x x\, f(x)$ jika X peubah acak diskrit ; dan

$E[X] = \int_{-\infty}^{\infty} f(x)\, dx$ jika X peubah acak kontinu

Teorema. Misalkan X suatu peubah acak dengan distribusi peluang f(x). Nilai harapan fungsi g(x) adalah :

$E[g(X)] = \sum_x g(x)\ f(x)$; jika X peubah acak diskrit ; dan

$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx$; jikaX peubah acak kontinu

Teorema. Variansi (keragaman) peubah acak X, ditulis σ^2 , adalah

$$\sigma^2 = E[X^2] - \mu^2$$

dengan : $E[X^2] = \sum_x x^2\, f(x)$; jika X peubah acak diskrit

$E[X^2] = \int_{-\infty}^{\infty} x^2\ f(x)\ dx$; jika X peubah acak kontinu

Contoh 1. Peubah Acak Diskrit

Dari percobaan lantunan sebuah koin sebanyak 3 kali. Jika X banyaknya angka yang muncul.

Hitung : a. μ_X b. σ_X^2

Penyelesaian. Tabel perhitungan ekspektasi diberikan sebagai berikut.

Tabel 3.4.1 Tabel perhitungan ekspektasi

x	0	1	2	3
f(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
x f(x)	0	$\frac{3}{8}$	$\frac{6}{8}$	$\frac{3}{8}$
x^2 f(x)	0	$\frac{3}{8}$	$\frac{12}{8}$	$\frac{9}{8}$

Maka : a. $\mu_X = E[X] = \sum_x x\, f(x) = 0 + \frac{3}{8} + \frac{6}{8} + \frac{3}{8} = \frac{12}{8} = \frac{3}{2}$

b. $\sigma_X^2 = E[X^2] - \mu_X^2 = (0 + \frac{3}{8} + \frac{12}{8} + \frac{9}{8}) - (\frac{3}{2})^2 = \frac{24}{8} - \frac{9}{4} = \frac{6}{8} = \frac{3}{4}$

Contoh 2. Peubah Acak Kontinu

PA kontinu X mempunyai fungsi padat peluang : $f(x) = \begin{cases} x^2/3 & , \ -1 < x < 2 \\ 0 & , \text{untuk x lain} \end{cases}$

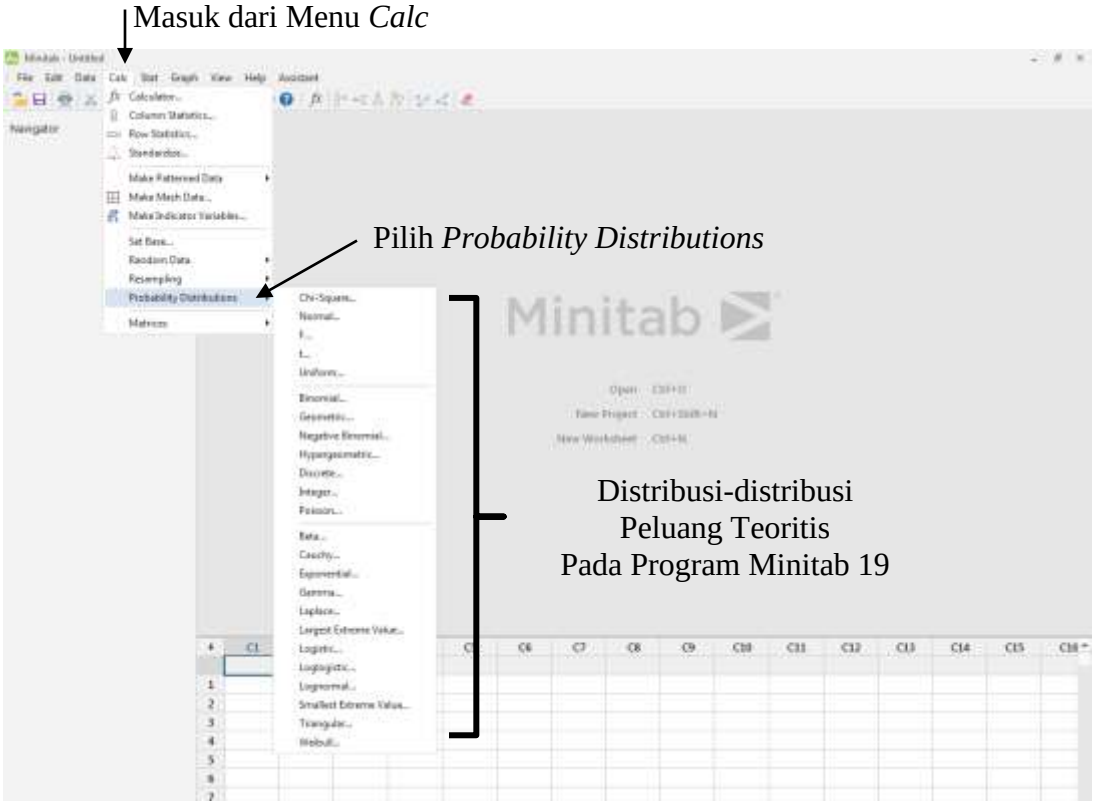
Maka:

$$E[X]=\int_{-1}^{2} x.\tfrac{1}{3}x^2dx=\tfrac{1}{12}x^4\Big|_{-1}^{2}=\frac{5}{4}$$

$$\sigma_X^2=E[X^2]-\{E[X]\}^2=\int_{-1}^{2} x^2.\tfrac{1}{3}x^2dx-\left(\frac{4}{5}\right)^2=\tfrac{1}{15}x^5\Big|_{-1}^{2}-\frac{16}{25}=\frac{11}{5}-\frac{16}{25}=\frac{39}{55}$$

3.5. Beberapa Distribusi Peluang Teoritis

Banyak peubah acak yang dihasilkan dalam suatu percobaan statistika mempunyai sifat yang sama dan pada dasarnya dapat dinyatakan dengan distribusi peluang yang sama, dengan demikian dapat dinyatakan dengan rumus yang sama. Berikut ini disajikan daftar distribusi peluang teoritis pada Program Minitab 19 (Gambar 3.5.1). Pada Gambar 3.5.1 ditampilkan cara untuk menggunakan *toolbar* atau ikon distribusi peluang (*probability distributions*). Dimulai dengan mengklik ikon *Calc* dan kemudian pilih ikon *Probability Distributions*. Maka akan terlihat semua distribusi peluang peubah acak yang ada pada program ini.



Gambar 3.5.1 Daftar distribusi peluang teoritis pada Program Minitab 19

3.5.1.Beberapa distribusi peluang peubah acak diskrit

a. Distribusi Seragam (*Uniform*)

Bila peubah acak X mendapat nilai x_1, x_2, ..., x_k, dengan peluang yang sama, maka distribusi seragam diskrit diberikan oleh,

$$f(x;k) = \frac{1}{k}\ ,\ x = x_1, x_2, ..., x_k$$

Contoh distribusi seragam adalah lantunan sebuah koin dan lantunan sebuah dadu.

- Dari lantunan sebuah koin : $f(x) = \frac{1}{2}$, x = sisi Gambar, sisi Angka
- Dari lantunan sebuah dadu : $f(x) = \frac{1}{6}$, x = 1, 2, 3, 4, 5, 6

b. Distribusi Binomial

Distribusi binomial didasarkan pada percobaan binomial, yang mempunyai 3 syarat berikut :

1. Percobaan terdiri atas n usaha yang berulang dan tiap usaha saling bebas.
2. Tiap usaha memberikan 2 hasil yang disebut dengan sukses dan gagal.
3. Peluang sukses disebut p, dan peluang gagal disebut q = 1 – p, tidak berubah dari usaha yang satu ke usaha berikutnya.

Bila suatu percobaan binomial dapat menghasilkan sukses dengan peluang p dan gagal dengan peluang q = 1 – p, maka distribusi peluang peubah acak binomial X, yaitu banyaknya sukses dalam n usaha bebas, ialah,

$$f(x) = P(X = x) = b(x;n,p) = \binom{n}{x} p^x (1-p)^{n-x}\ ,\ x = 0, 1, 2, ..., n$$

dengan rataan $\mu = np$ dan $\sigma^2 = npq$.

Catatan. Berikut catatan penting untuk hitung peluang distribusi Binomial dan yang berlaku umum untuk semua peubah acak diskrit.

1. Menurut Hukum Peluang : $\sum_{x=0}^{n} f(x) = f(0) + f(1) + ... + f(n) = \sum_{x=0}^{n} b(x;n,p) = 1$

   Yaitu jumlah nilai peluang untuk semua x = 0, 1, 2, .., n adalah 1

   (Ingat contoh sederhana: koin, ada 2 titik sampel, masing-masing peluang ½, jumlahnya 1)
2. Untuk menghitung peluang distribusi binomial digunakan Tabel Peluang Binomial yang dapat kita buat sendiri dengan menggunakan Program Minitab.

Ada 2 jenis tabel yaitu :

a. Tabel Distribusi Peluang.

Nilai dalam Tabel Distribusi Peluang adalah $P(X = x) = f(x) = b(x; n, p)$.

b. Tabel Distribusi Kumulatif.

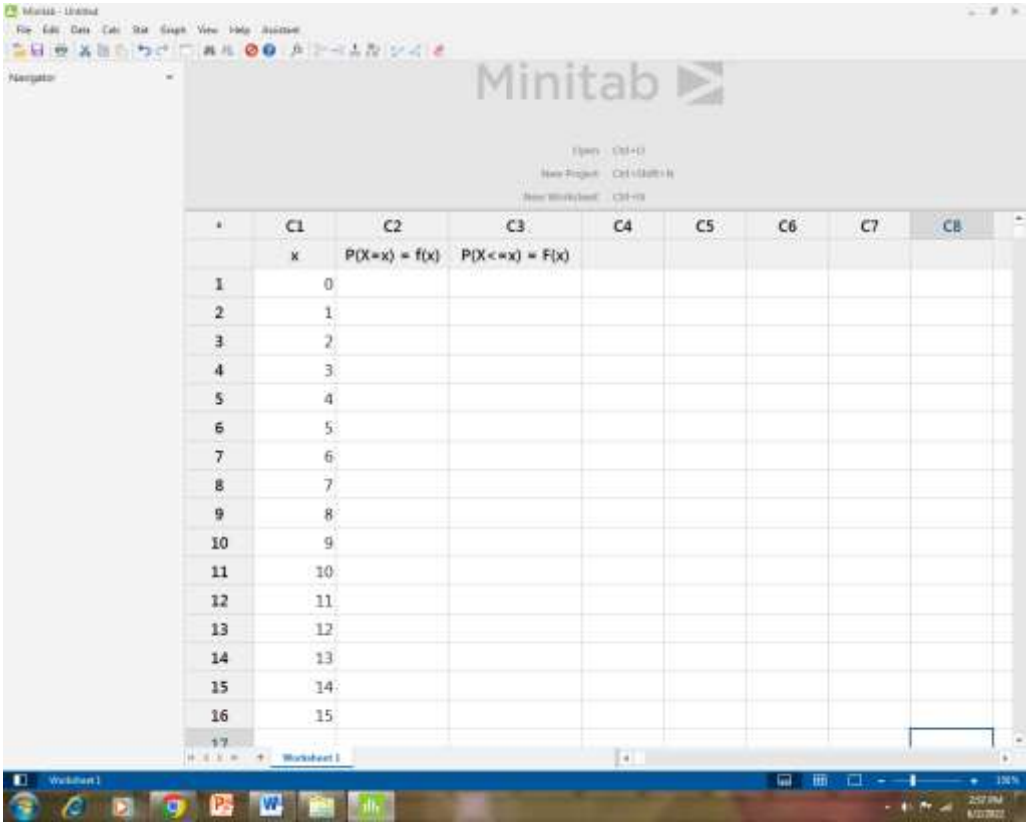
Nilai dalam Tabel Distribusi Kumulatif adalah $P(X \leq r)$ yaitu :

$$P(X \leq r) = F(r) = P(X = 0) + P(X = 1) + P(X = 2) + \ldots + P(X = r - 1) + P(X = r)$$

$$= \sum_{x=0}^{r} b(x; n, p) \quad \rightarrow \text{disebut peluang kumulatif (}\textit{cumulative probability}\text{)}$$

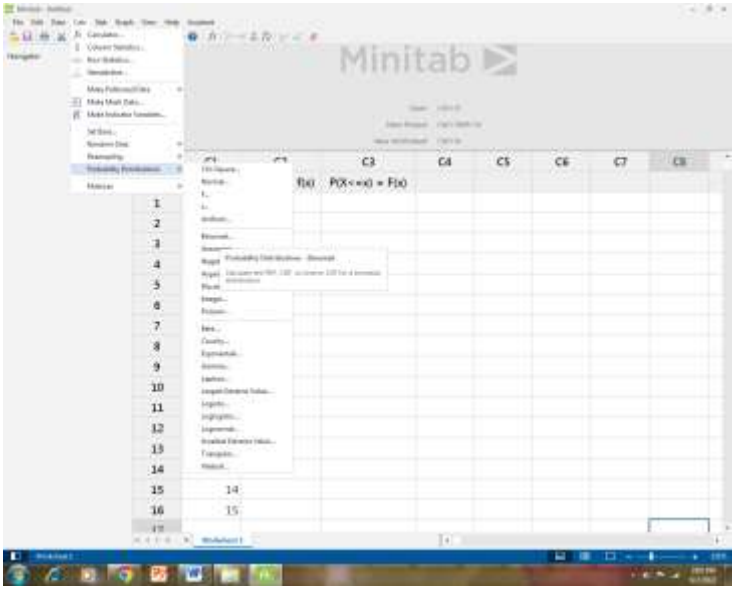
Berikut langkah membuat Tabel Distribusi Peluang f(x) dan Tabel Distribusi Kumulatif F(x) dengan menggunakan Program Minitab 19, untuk n = 15 dan p = 0,6.

Langkah 1. Karena n = 15 maka nilai x = 0, 1, 2, …, 15. Isikan nilai x pada kolom C1 seperti pada Gambar 3.5.1.1 berikut. Pada kolom C2 persiapkan kolom untuk nilai P(X = x) = f(x) dan pada kolom C3 persiapkan kolom untuk nilai P(X <= x) = F(x). Hati-hati, perhatikan penulisan "f".



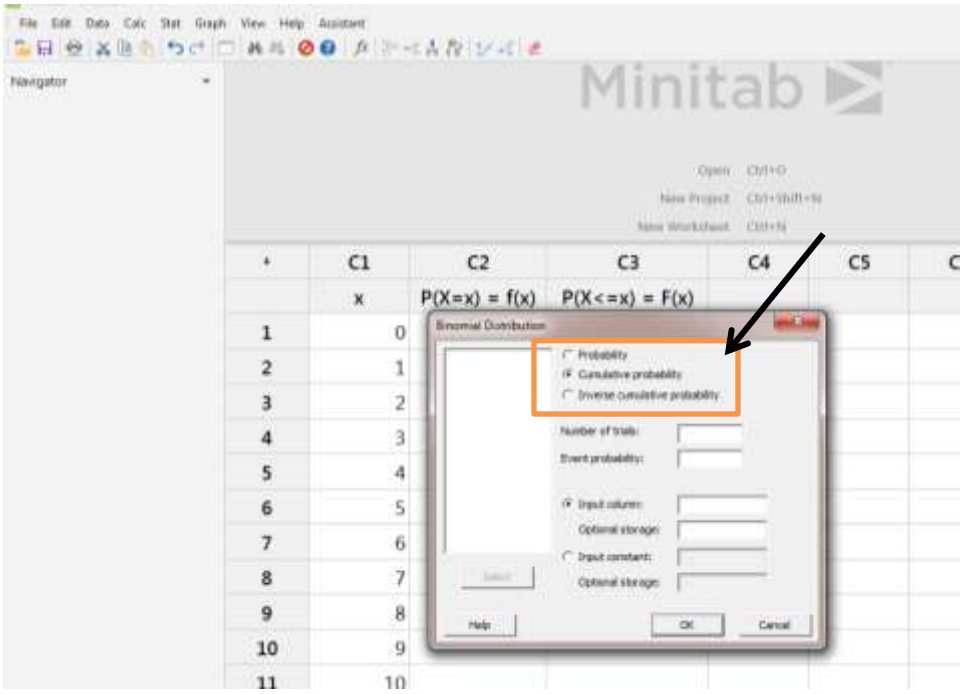
Gambar 3.5.1.1 Langkah 1 menu Distribusi Binomial

Langkah 2. Klik ikon *Calc*, kemudian ikon *Probability Distributions* dan terakhir ikon Binomial seperti pada Gambar 3.5.1.2 berikut ini.



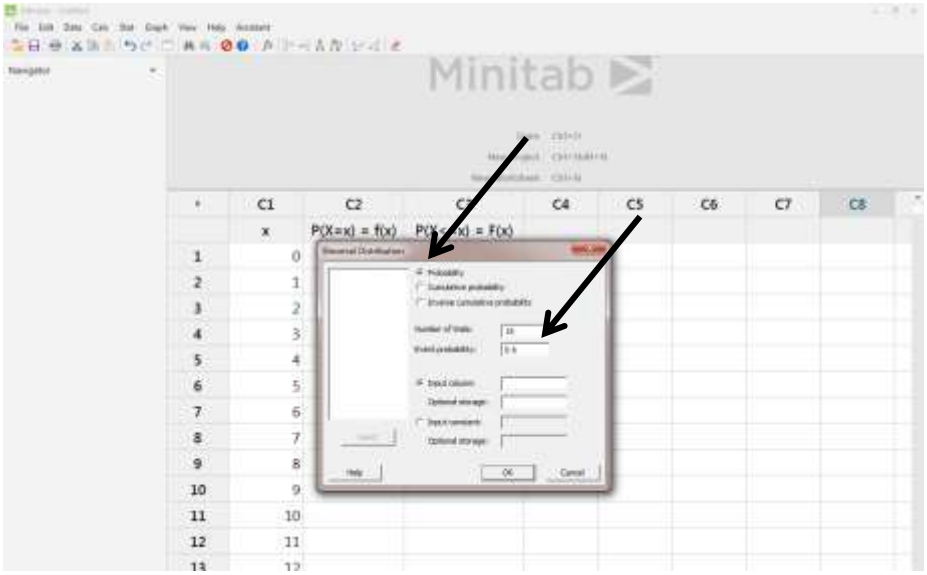
Gambar 3.5.1.2 Langkah 2 menu Distribusi Binomial

Maka akan tampil seperti pada Gambar 3.5.1.3 berikut ini. Ada 3 pilihan yaitu *Probability* f(x), *Cumulative probability* F(x) dan *Inverse cumulative probability*. *Default*nya adalah *Cumulative probability* F(x).



Gambar 3.5.1.3 Pilihan menu Distribusi Binomial

Langkah 2. Kita akan menentukan nilai pada C2 yaitu *Probability* f(x), maka pindahkan kursor (titik hitam pada *Cumulative probability*) ke *Probability* seperti pada Gambar 3.5.1.4 berikut. Kemudian isikan nilai-nilai seperti pada Gambar 3.5.1.4 yaitu *Number of trials* 15 dan *Event probability* 0,6 seperti yang ditunjukkan oleh tanda panah pada Gambar 3.5.1.4.



Gambar 3.5.1.4 Input data *probability*

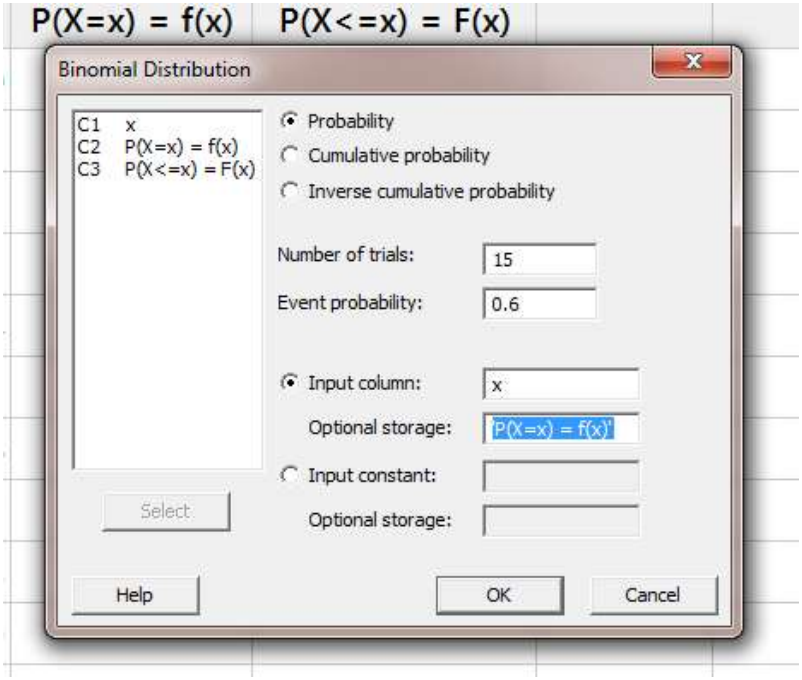
Langkah 3. Tempatkan kursor pada *Input column*, maka pada kotak kosong di sebelah kiri akan muncul C1 x, yaitu data x yang akan kita tentukan nilai peluangnya atau f(x) nya. Kemudian arahkan kursor ke C1 x dan klik sehingga berwarna biru. Lalu klik Select yang terletak di bawah sebelah kiri seperti yang ditunjukkan oleh tanda panah (Gambar 3.5.1.5).

Gambar 3.5.1.5 Langkah 3 menu Distribusi Binomial

Maka akan tampil seperti pada Gambar 3.5.1.6 berikut ini. Sekarang di sebelah kiri telah masuk C2 P(X=x) = f(x) dan C3 P(X<=x) = F(x). Karena kita hendak menentukan P(X=x) = f(x), arahkan kursor ke C2 P(X=x) = f(x) sehingga berwarna biru, kemudian klik *Select*.

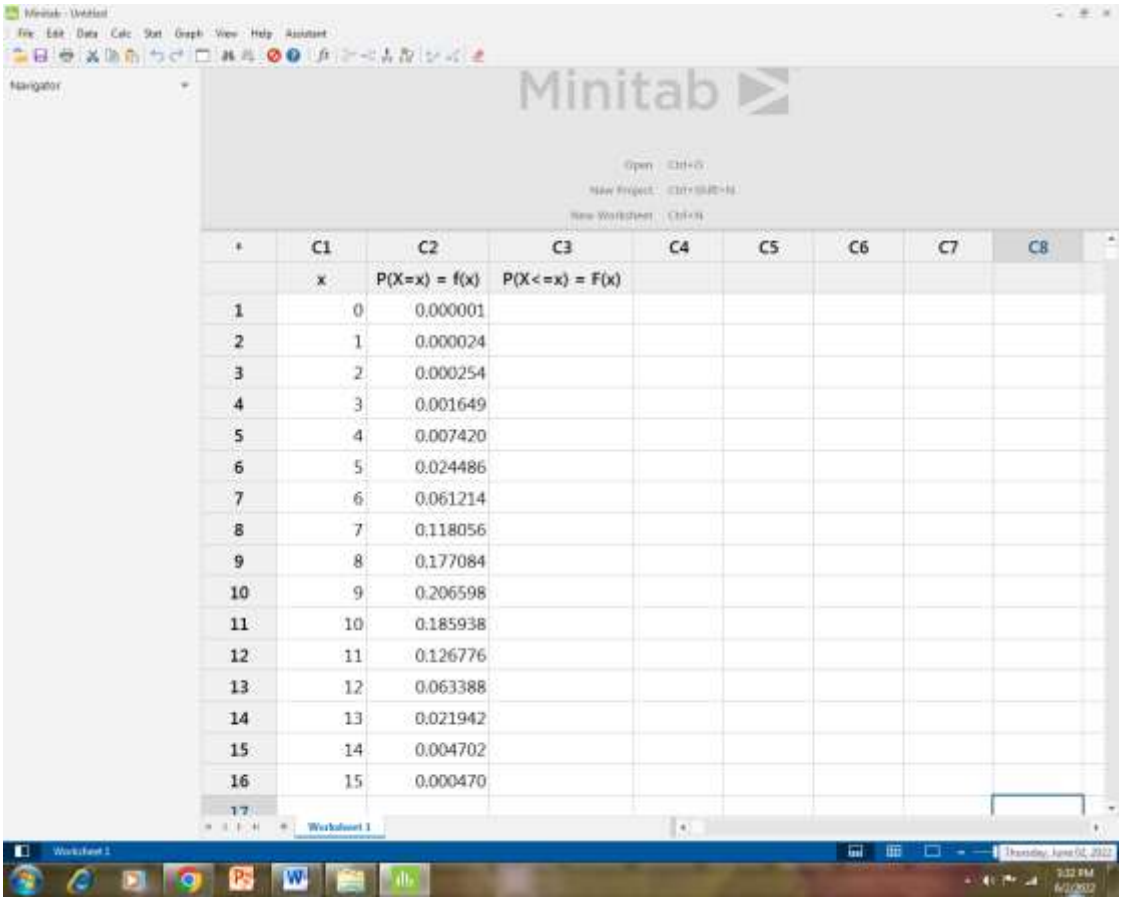
Gambar 3.5.1.6 Input data dari kolom pada menu *Probability*

Maka C2 P(X=x) = f(x) akan masuk ke *Optional storage* yang bermakna nilai P(X=x) = f(x) akan disimpan pada kolom C2 seperti yang ditunjukkan pada Gambar 3.5.1.7



Gambar 3.5.1.7 Menu *Optional storage*

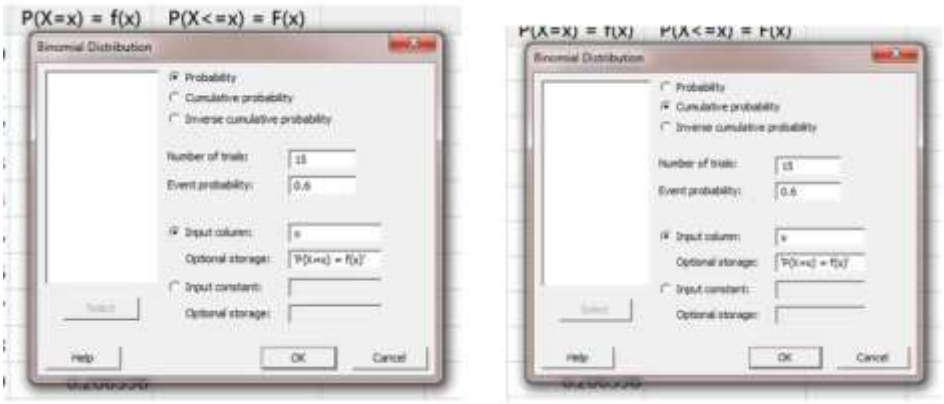
Selanjutnya klik Ok. Maka pada kolom C2 akan terisi nilai-nilai P(X=x) = f(x) untuk semua x, yaitu x = 0, 1, 2, …, 15, seperti diperlihatkan pada Gambar 3.5.1.8



	C1	C2	C3	C4	C5	C6	C7	C8
	x	P(X=x) = f(x)	P(X<=x) = F(x)					
1	0	0.000001						
2	1	0.000024						
3	2	0.000254						
4	3	0.001649						
5	4	0.007420						
6	5	0.024486						
7	6	0.061214						
8	7	0.118056						
9	8	0.177084						
10	9	0.206598						
11	10	0.185938						
12	11	0.126776						
13	12	0.063388						
14	13	0.021942						
15	14	0.004702						
16	15	0.000470						

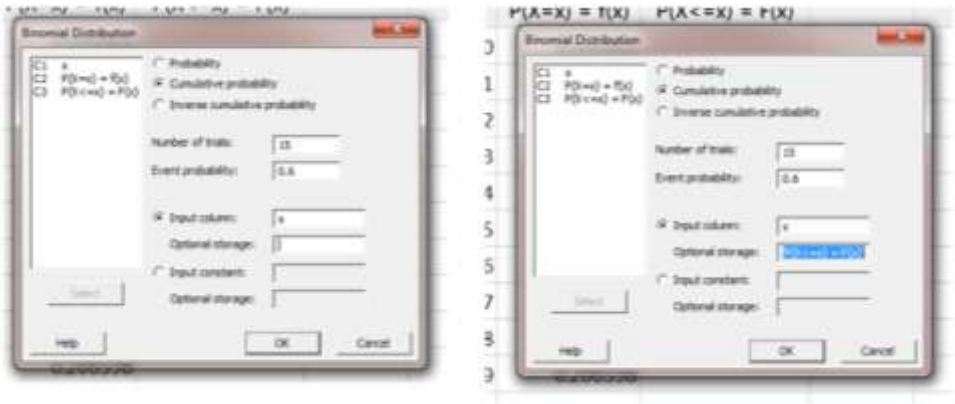
Gambar 3.5.1.8 *Output Probability* Binomial

Langkah 4. Sekarang kita akan menentukan Tabel Distribusi Kumulatif F(x). Jika kita mulai kembali dari ikon *Calc*, *Probability distributions* dan Binomial; maka akan tampil seperti pada Gambar 3.5.1.9 berikut ini. Pindahkan kursor ke *Cumulative distribution*.



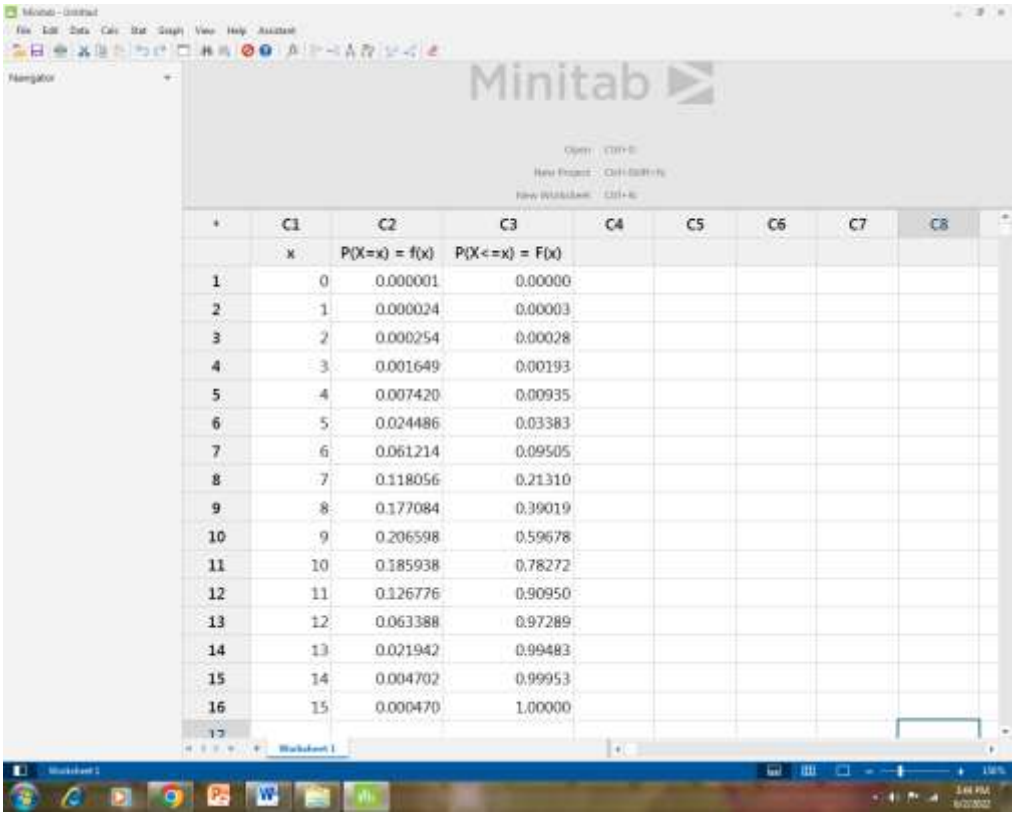
Gambar 3.5.1.9 Langkah 4 menu Distribusi Binomial

Langkah 5. Hapus P(X = x) = f(x) dari *Optional Storage*. Dan klik C3 P(X<=x) = F(x) sehingga berwarna biru lalu klik *Select*, dan terakhir Ok (Gambar 3.5.1.10).



Gambar 3.5.1.10 Langkah 5 menu Distribusi Binomial

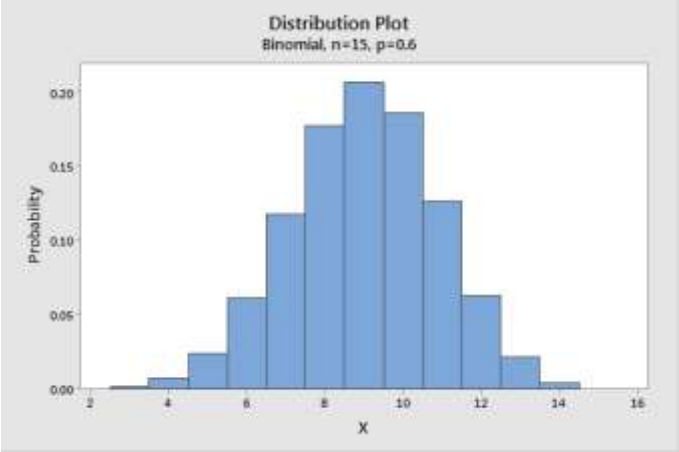
Maka nilai P(X<=x) = F(x) akan diisikan pada C3 P(X<=x) = F(x) seperti yang diperlihatkan pada Gambar 3.5.1.11 berikut.



	C1	C2	C3	C4	C5	C6	C7	C8
	x	P(X=x) = f(x)	P(X<=x) = F(x)					
1	0	0.000001	0.00000					
2	1	0.000024	0.00003					
3	2	0.000254	0.00028					
4	3	0.001649	0.00193					
5	4	0.007420	0.00935					
6	5	0.024486	0.03383					
7	6	0.061214	0.09505					
8	7	0.118056	0.21310					
9	8	0.177084	0.39019					
10	9	0.206598	0.59678					
11	10	0.185938	0.78272					
12	11	0.126776	0.90950					
13	12	0.063388	0.97289					
14	13	0.021942	0.99483					
15	14	0.004702	0.99953					
16	15	0.000470	1.00000					

Gambar 3.5.1.11 *Output Cumulative probability* Binomial

Grafik distribusi peluang peubah acak Binomial dengan n = 15 dan p = 0,6 diberikan pada Gambar 3.5.1.12 berikut ini.



Gambar 3.5.1.12 Grafik distribusi peluang Binomial, n = 15 dan p = 0,6

Berikut ini hitung peluang yang sering dilakukan pada distribusi peubah acak diskrit seperti Binomial, yang perhitungannya dilakukan menggunakan peluang kumulatif F(x).

$$P(X < r) = P(X = 0) + P(X = 1) + \ldots + P(X = r - 1) = \sum_{x=0}^{r-1} b(x;n,p) = F(r-1)$$

$$P(X \geq r) = 1 - P(X < r) = 1 - \sum_{x=0}^{r-1} b(x;n,p) = 1 - F(r-1)$$

$$P(X > r) = 1 - P(X \leq r) = 1 - \sum_{x=0}^{r} b(x;n,p)$$

$$P(X = r) = f(r) = P(X \leq r) - P(X < r - 1) = \sum_{x=0}^{r} b(x;n,p) - \sum_{x=0}^{r-1} b(x;n,p)$$

$$P(a \leq X \leq b) = P(X = a) + P(X = a + 1) + P(X = a + 2) + \ldots + P(X = b)$$

$$= P(X \leq b) - P(X < a - 1) = \sum_{x=0}^{b} b(x;n,p) - \sum_{x=0}^{a-1} b(x;n,p)$$

$$P(a < X \leq b) = P(X = a + 1) + P(X = a + 2) + \ldots + P(X = b)$$

$$= P(X \leq b) - P(X \leq a) = \sum_{x=0}^{b} b(x;n,p) - \sum_{x=0}^{a} b(x;n,p)$$

Contoh. Suatu peubah acak binomial mempunyai n = 15 dan p = 0,6. Hitung peluang-peluang berikut : a. $P(X = 5)$ b. $P(X \leq 4)$ c. $P(3 \leq X \leq 7)$ d. $P(X \geq 8)$

Penyelesaian

Berikut adalah Tabel Binomial untuk n = 15 dan p = 0,6 yang telah kita kerjakan sebelumnya.

Tabel 3.5.1.1 Tabel Binomial untuk n = 15 dan p = 0,6

x	P(X=x) = f(x)	P(X<=x) = F(x)
(1)	(2)	(3)
0	0.000001	0.00000
1	0.000024	0.00003
2	0.000254	0.00028
3	0.001649	0.00193
4	0.007420	0.00935
5	0.024486	0.03383
6	0.061214	0.09505
7	0.118056	0.21310
8	0.177084	0.39019
9	0.206598	0.59678
10	0.185938	0.78272
11	0.126776	0.90950
12	0.063388	0.97289
13	0.021942	0.99483
14	0.004702	0.99953
15	0.000470	1.00000

a. P(X = 5). Dengan menggunakan Tabel Distribusi Peluang di atas :

(i) Menggunakan kolom 2: P(X = 5) = 0,024486

(ii) Menggunakan kolom 3:

$$P(X = 5) = \sum_{x=0}^{5} b(x;10,0.25) - \sum_{x=0}^{4} b(x;10,0.25) = 0{,}03383 - 0{,}00935 = 0{,}2448$$

b. P(X ≤ 4). Dengan menggunakan kolom 3 untuk x = 4 : $P(X \le 4) = \sum_{x=0}^{4} b(x;10,0.25) = 0{,}00935$

c. $P(3 \le X \le 7) = \sum_{x=0}^{7} b(x;10,0.25) - \sum_{x=0}^{2} b(x;10,0.25) = 0{,}21310 - 0{,}00028 = 0{,}21282$

d. $P(X \ge 8) = 1 - \sum_{x=0}^{7} b(x;10,0.25) = 1 - 0{,}21310 = 0{,}7869$

c. Distribusi Hipergeometrik

Distribusi hipergeometrik didasarkan pada suatu percobaan Hipergeometrik yang memiliki sifat sebagai berikut :

1. sampel acak ukuran n diambil dari N benda
2. sebanyak M benda disebut sukses, sedangkan sisanya, N – M, disebut gagal

Distribusi peluang peubah acak hipergeometrik X, yaitu banyaknya sukses dalam sampel acak ukuran n yang diambil dari N benda yang mengandung M sukses dan N – M gagal, ialah,

$$P(X=x)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},\ x=0,1,2,.....,n$$

dengan rataan dan variansi masing-masing diberikan oleh

$$\mu=\frac{nM}{N} \qquad \text{dan} \qquad \sigma^2=\frac{N-n}{N-1}.n.\frac{M}{n}.\left(1-\frac{M}{n}\right)$$

Contoh. Tiga orang mahasiswa ITI akan dipilih untuk mengikuti sebuah seminar. Mahasiswa yang memenuhi kriteria untuk dipilih terdiri dari 4 orang mahasiswa Teknik Mesin dan 3 orang mahasiswa Teknik Industri. Buatlah rumus dan tabel distrbusi peluang dari banyaknya mahasiswa Teknik Mesin yang terpilih ikut seminar.

Penyelesaian

Misalkan X menyatakan banyaknya mahasiswa Teknik Mesin yang terpilih ikut seminar, maka dalam 3 mahasiswa yang akan dipilih, nilai-nilai X dapat bernilai : 0, 1, 2, atau 3. Dan rumus distribusi peluangnya adalah

$$P(X = x) = f(x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}},\ x = 0, 1, 2, 3$$

Untuk membuat Tabel Distribusi Peluang dapat dilakukan dengan dua cara:

a. Kita hitung terlebih dahulu nilai-nilai f(x) untuk setiap x , dan yang pertama dihitung adalah kombinasi 7 dan 3 yaitu :

$$\binom{7}{3}=\frac{7!}{3!\,(7-3)!}=\frac{7!}{3!\,4!}=\frac{7.6.5.4!}{3.2.1.4!}=35$$

Maka :

$$f(0)=\frac{\binom{4}{0}\binom{3}{3}}{\binom{7}{3}}=\frac{(1).(1)}{35}=\frac{1}{35}= \qquad f(1)=\frac{\binom{4}{1}\binom{3}{2}}{\binom{7}{3}}=\frac{(4).(3)}{35}=\frac{12}{35}$$

$$f(2) = \frac{\binom{4}{2}\binom{3}{1}}{\binom{7}{3}} = \frac{(6).(3)}{35} = \frac{18}{35} \qquad f(3) = \frac{\binom{4}{3}\binom{3}{0}}{\binom{7}{3}} = \frac{(4).(1)}{35} = \frac{4}{35}$$

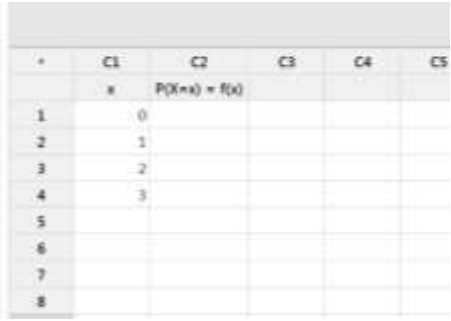
Dan Tabel Distribusi Peluangnya adalah sebagai berikut.

Tabel 3.5.1.2 Tabel hipergeometrik N = 7, n = 3, M = 3

x	0	1	2	3
f(x)	1/35	12/35	18/35	4/35

b. Dengan menggunakan Program Minitab.

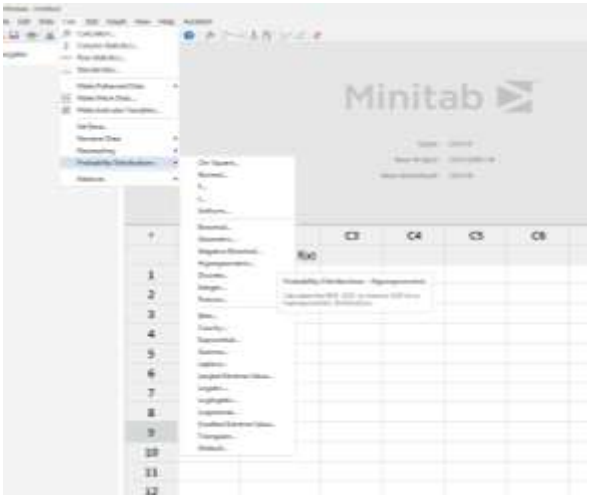
Isikan nilai x = 0, 1, 2 dan 3 pada kolom C1 dan P(X=x) = f(x) pada kolom C2 seperti pada Gambar 3.5.1.13 berikut ini.



Gambar 3.5.1.13 Input data Hipergeometrik

Langkahnya sama seperti menentukan nilai P(X = x) pada Distribusi Binomial.

Di sini dari ikon *Probability Distribution*, pilih *Hypergeometric* (Gambar 3.5.1.14).



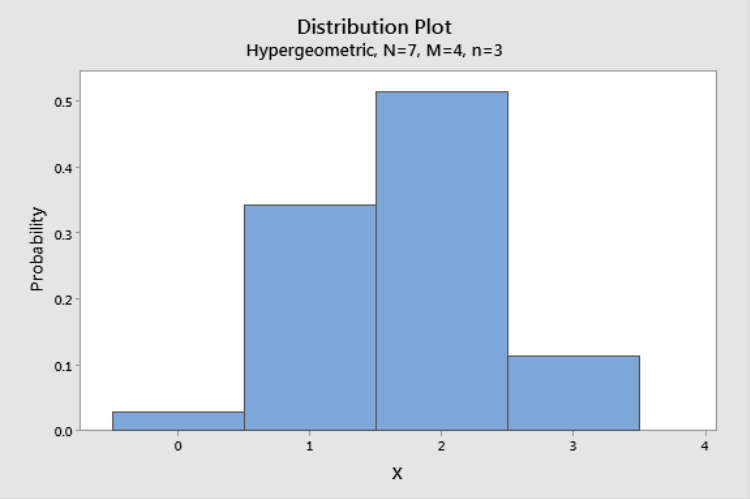
Gambar 3.5.1.14 Langkah pertama menu Distribusi Hipergeometrik

Lalu pindahkan kursor dari *Cumulative probability* ke *Probability*. Kemudian isikan nilai-nilai pada *Population size* (N), *Event count in population* (M) dan *Sample size* (n). Lalu letakkan kursor di *Input column* dan seterusnya seperti pada Distribusi Binomial. Kemudian klik Ok. Hasilnya akan diisikan ke dalam C2 P(X = x) = f(x) (Gambar 3.5.1.15) yang sama dengan nilai peluang pada Tabel 3.5.1.2.

	C1	C2
	x	P(X=x) = f(x)
1	0	0.028571
2	1	0.342857
3	2	0.514286
4	3	0.114286
5		

Gambar 3.5.1.15 *Output Probability* Hipergeometrik

Grafik distribusi peluang Hipergeometrik untuk contoh di atas diberikan pada Gambar 3.5.1.16 berikut.



Gambar 3.5.1.16 Grafik distribusi peluang hipergeometrik N = 7, n = 3, M = 3

d. Distribusi Poisson

Distribusi peluang peubah acak Poisson X, yang menyatakan banyaknya sukses yang terjadi dalam suatu selang waktu atau daerah tertentu, diberikan oleh,

$$P(X=x) = \frac{e^{-\mu}\,\mu^{x}}{x!}\ ,\ x = 0,1,2,.....$$

dengan μ menyatakan rata-rata banyaknya sukses yang terjadi dalam selang waktu atau daerah tertentu tersebut dan e = 2,71828.... . Rataan dan variansi distribusi Poisson adalah μ.

Sama halnya seperti peubah acak Binomial dan Hipergeometrik, penentuan nilai peluang peubah acak Poisson dapat dilakukan dengan menggunakan Program Minitab.

Contoh.

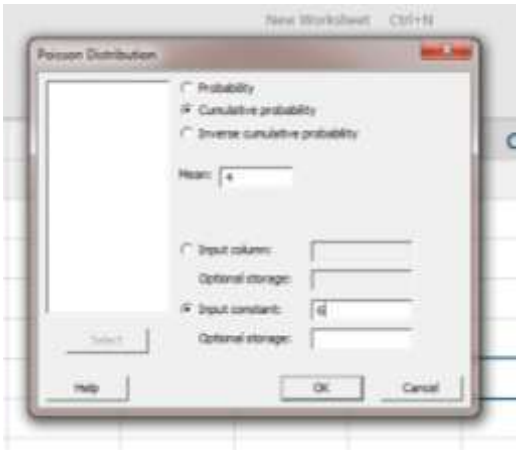
Rata-rata banyaknya partikel radioaktif yang melewati suatu penghitung selama 1 milidetik dalam suatu percobaan di laboratorium adalah empat. Berapakah peluang enam partikel melewati penghitung dalam suatu milidetik tertentu?

Penyelesaian $\mu = 4$ dan yang ditanyakan : P(X = 6)

Berikut ini langkah menghitung peluang P(X = 6) pada distribusi Poison dengan menggunakan Program Minitab. Seperti langkah pada Distribusi Binomial, dari Menu *Stat* pilih ikon *Probability Distributions* dan kemudian pilih Poisson. Maka akan muncul kotak dialog *Poisson Distribution* seperti pada Gambar 3.5.1.17.

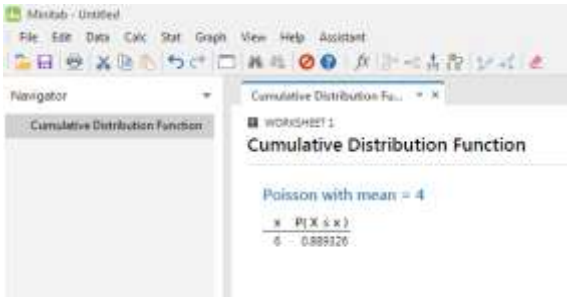
Gambar 3.5.1.17 Kotak dialog *Poisson Distribution*

Kemudian pada kotak dialog Poisson *Distribution* pindahkan kursor ke *Probability*, lalu isikan nilai *mean* 4. Selanjutnya letakkan kursor pada *Input constant* dan isikan nilanya 6 (Gambar 3.5.1.18).



Gambar 3.5.1.18 Input data Poison

Kemudian klik Ok. Hasilnya diberikan seperti pada Gambar 3.5.1.19 berikut.



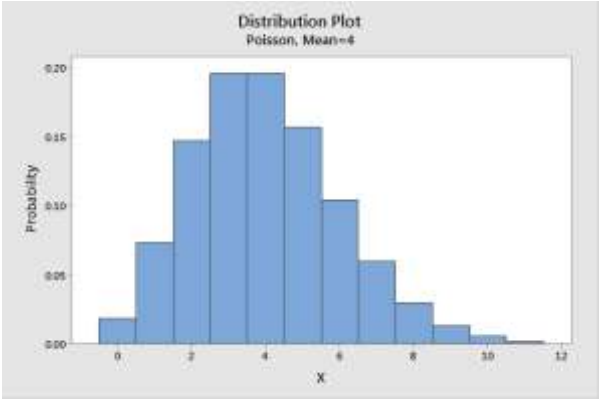
Gambar 3.5.1.19 *Output Probability* Poison

Yaitu: *Poisson with mean* = 4

x	P(X ≤ x)
6	0.889326

yang memberikan nilai P(X = 6) = 0,889326

Grafik distribusi peluang Poisson untuk contoh di atas diberikan pada Gambar 3.5.1.20 berikut.

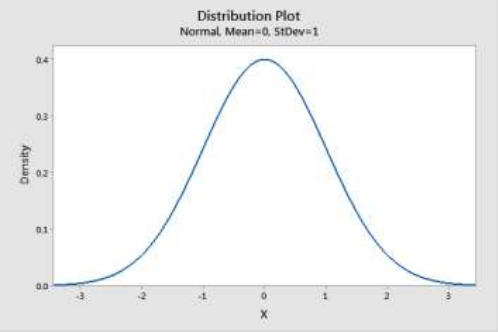


Gambar 3.5.1.20 Grafik distribusi peluang Poison μ = 4

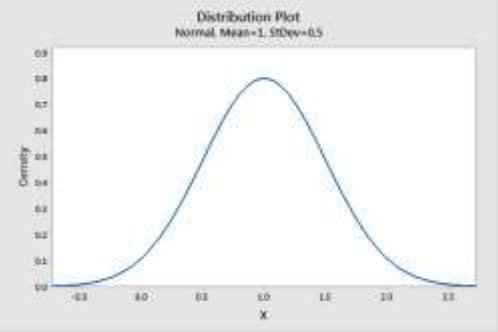
3.5.2. Beberapa Distribusi Peluang Peubah Acak Kontinu

a. Distribusi Normal.

Distribusi peluang kontinu yang terpenting dalam statistika adalah distribusi normal. Grafiknya disebut kurva normal, berbentuk lonceng dan simetris terhadap rataannya (μ) seperti yang diperlihatkan pada Gambar 3.5.2.1 dan Gambar 3.5.2.2 di bawah ini, menggambarkan berbagai kumpulan data yang muncul di alam, industri dan penelitian.



Gambar 3.5.2.1 Kurva Normal $\mu = 0$, $\sigma^2 = 1$



Gambar 3.5.2.2 Kurva Normal $\mu = 1$, $\sigma^2 = 0{,}25$

Sebagaimana telah disebutkan dalam konsep distribusi peluang peubah acak kontinu, maka pada Distribusi Normal P(X = x) adalah nol, sebab nilai-nilai peubah acak kontinu terdefinisi pada suatu selang atau interval sehingga nilai peluangnya terdefinisi pada selang-selang tersebut. Maka dalam penggunaan Program Minitab kita akan menggunakan dua pilihan saja yaitu *Cumulative probability* dan *Inverse cumulative probability*.

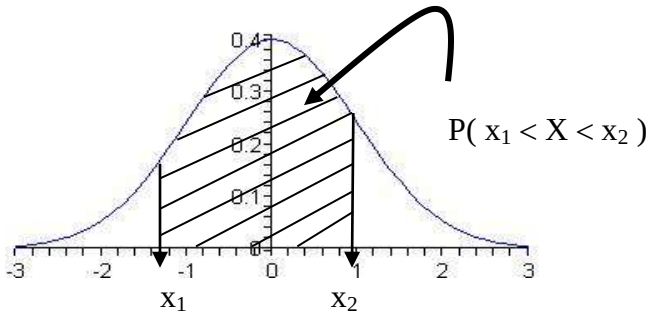
Fungsi padat peluang (*probability density function*) peubah acak normal X, dengan rataan μ dan variansi σ^2, ialah :

$$n(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}.\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],\ -\infty<x<\infty$$

Perhitungan nilai peluang pada suatu selang untuk peubah acak kontinu dihitung menggunakan Kalkulus Integral yang tafsirannya sebagai luas daerah di bawah kurva normal pada selang tersebut, dengan luas seluruhnya adalah 1 (satu). Maka $P(x_1 < X < x)$ adalah:

$$P(x_1<X<x_2)=\int_{x1}^{x2} n(x;\mu,\sigma)=\int_{x1}^{x2}\ \frac{1}{\sigma\sqrt{2\pi}}.\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\,dx$$

seperti yang diperlihatkan pada Gambar 3.5.2.3 berikut bahwa P($x_1 < X < x_2$) dinyatakan dengan luas daerah yang diarsir.



Gambar 3.5.2.3 P($x_1 < X < x_2$) Distribusi Normal

Peubah acak distribusi normal yang banyak digunakan adalah Distribusi Normal Baku (peubah acaknya diberi notasi Z) yang didefinisikan sebagai berikut.

Definisi. Distribusi peubah acak normal dengan $\mu = 0$ dan $\sigma^2 = 1$ disebut peubah acak Z berdistribusi normal baku. Grafik Distribusi Normal Baku diberikan pada Gambar 3.5.2.1 di atas.

Berikut ini akan dibahas hitung peluang peubah acak Distribusi Normal dengan Program Minitab menggunakan pilihan *Cumulative probability*, yang akan berlaku umum untuk semua distribusi peubah acak kontinu.

1. $P(X < x)$
2. $P(x_1 < X < x_2) = P(X < x_2) - P(X < x_1)$
3. $P(X > x) = 1 - P(X < x)$

Contoh.

Suatu perusahaan listrik menghasilkan bola lampu yang umurnya berdistribusi normal dengan rataan 800 jam dan simpangan baku 40 jam.

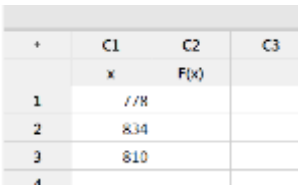
a. Hitung peluang suatu bola lampu dapat menyala antara 778 dan 834 jam.
b. Hitung peluang bola lampu dapat menyala lebih dari 810 jam.

Penyelesaian

Kita akan menggunakan Program Minitab untuk menentukan peluang kumulatif F(x) untuk :

x = 778, 834 dan 810 jam.

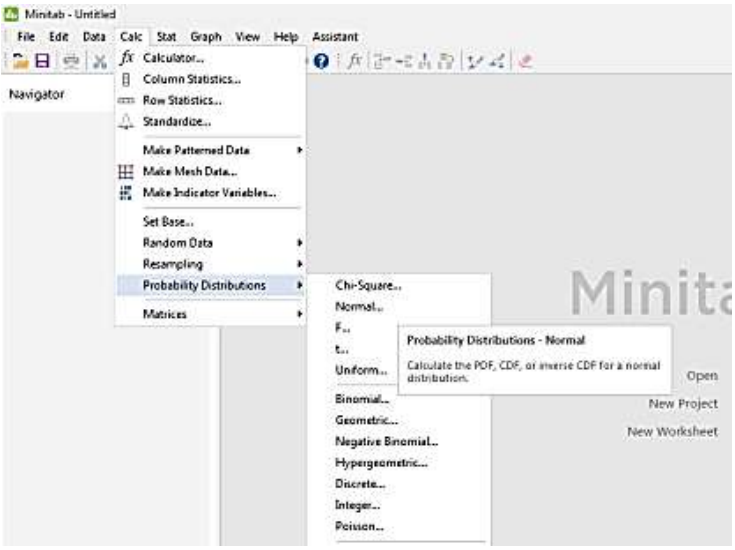
Maka siapkan tempat untuk nilai x dan F(x) pada *worksheet* seperti pada Gambar 3.5.2.4 berikut. Langkah ini sama saja seperti langkah pada Distribusi Binomial yang telah dibahas sebelumnya.



	C1	C2	C3
	x	F(x)	
1	778		
2	834		
3	810		

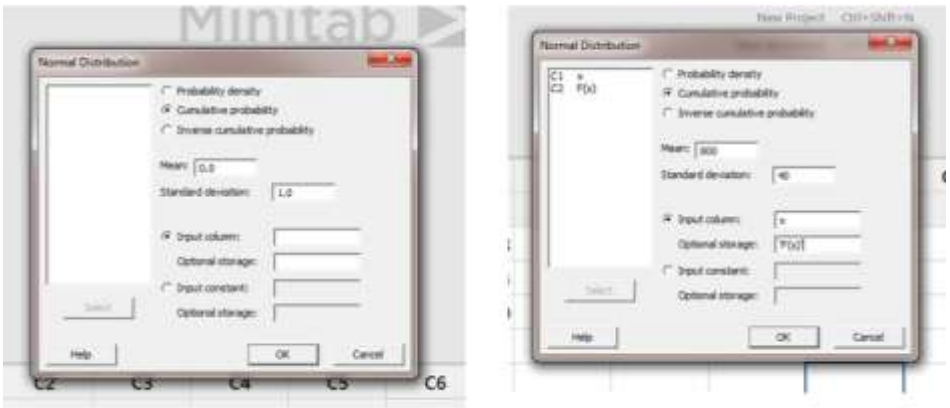
Gambar 3.5.2.4 Input data Normal

Kemudian dari ikon *Probability Distributions* pilih Normal.



Gambar 3.5.2.5 Menu *Probability Distributions* dan Normal

Akan tampil kotak dialog *Normal Distribution*. Kita akan mengisi nilai *Mean* dan *Standard Deviation*. *Default*nya adalah *Mean* = 0 dan *Standard Deviation* = 1. Ganti *Mean* = 800 dan *Standard Deviation* = 40. Selanjutnya lakukan seperti pada Distribusi Binomial: *Input column* C1 x, *Optional storage* C2 F(x) (Gambar 3.5.2.6).



Gambar 3.5.2.6 Input data Normal

Kemudian klik Ok. Maka nilai F(x) akan diisikan pada kolom C2 F(x) seperti pada Gambar 3.5.2.7 berikut.

	C1	C2	C3
	x	F(x)	
1	778	0.291160	
2	834	0.802337	
3	810	0.598706	
4			

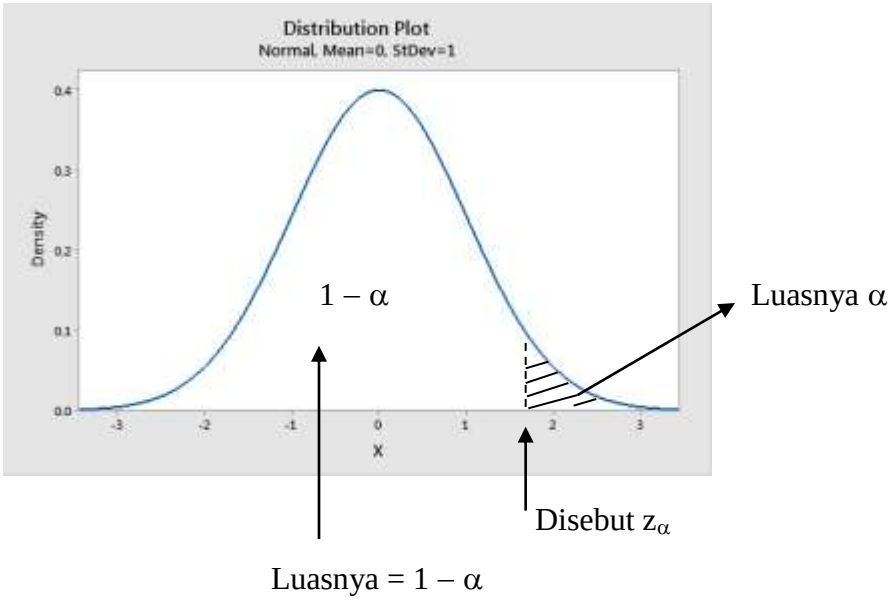
Gambar 3.5.2.7 *Output* Normal $\mu = 800$, $\sigma = 40$

Maka:

a. $P(778 < X < 834) = P(X < 834) - P(X < 778) = 0{,}802337 - 0{,}291160 = 0{,}511217$

b. $P(X > 810) = 1 - P(X < 810) = 1 - 0{,}598706 = 0{,}401294$

Distribusi Normal merupakan distribusi peubah acak kontinu yang banyak digunakan dalam Statistika Inferensi yaitu penyusunan selang kepercayaan dan pengujian hipotesis terhadap parameter populasi. Di sini konsep distribusi normal yang banyak digunakan adalah penentuan nilai z_α (dikenal dengan istilah z-*score*), yaitu nilai z pada distribusi normal baku Z sehingga luas di sebelah kanannya sebesar α, yang akan dijelaskan melalui Gambar 3.5.2.8 berikut.



Gambar 3.5.2.8 Ilustrasi z-*score*

Keterangan Gambar.

Jika luas yang di sebelah kanan z_α adalah α (yang diarsir) yaitu $P(Z > z_\alpha) = \alpha$,

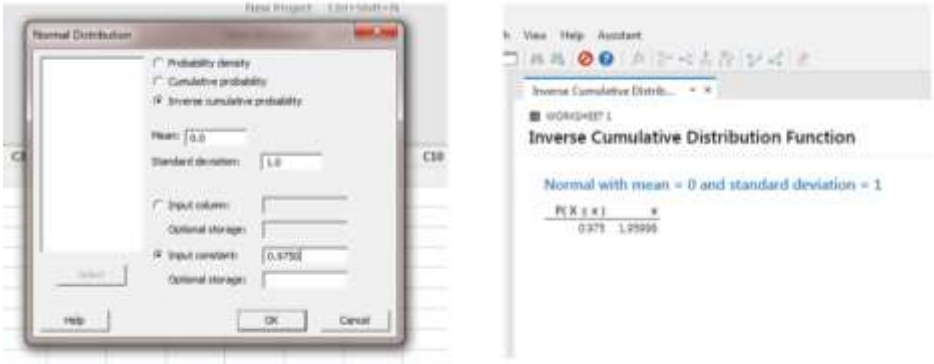
Maka luas yang di sebelah kiri z_α adalah $1 - \alpha$

yang dengan menggunakan peluang kumulatif dituliskan $P(Z < z_\alpha) = 1 - \alpha$

Jadi untuk menentukan nilai z_α digunakan peluang kumulatif: $P(Z < z_\alpha) = 1 - \alpha$ dan dapat ditentukan dengan *Inverse cumulative probability* pada Program Minitab. z-*score* yang paling sering dipakai adalah $z_{0,025}$, yang nilainya diperoleh sebagai berikut.

$$P(Z < z_{0,025}) = 1 - 0,025 = 0,975$$

Berikut ini diberikan cara menentukan nilai $z_{0,025}$ dengan menggunakan Program Minitab. Pilih menu *Inverse cumulative probability* ; dalam *Input constant* diisi $1 - \alpha = 1 - 0,0250 = 0,9750$ (Gambar 3.5.2.9).



Gambar 3.5.2.9 *Inverse cumulative probability* $z_{0,025}$

Maka diperoleh bahwa $z_{0,025} = 1,95996 \approx 1,96$.

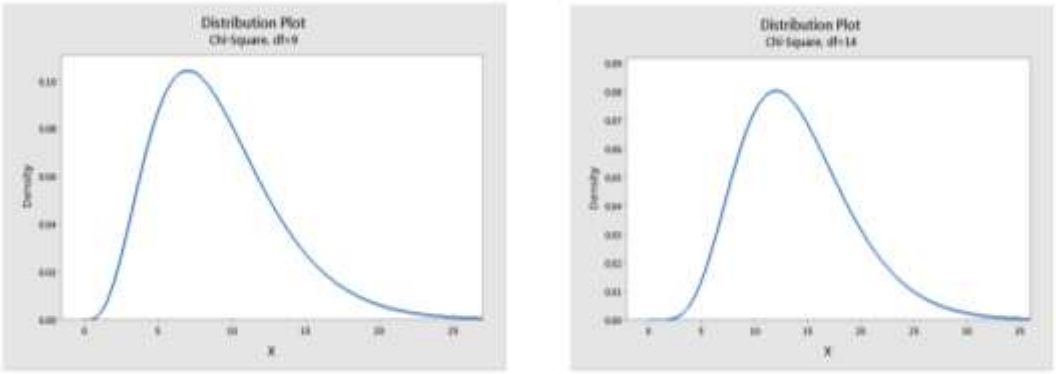
b. Distribusi Khi-kuadrat (*Chi-square*)

Sama hal nya dengan Distribusi Normal, Distribusi Khi-kuadrat juga merupakan salah satu distribusi peubah acak kontinu yang banyak digunakan dalam Statistika Inferensi khususnya yang berkaitan dengan parameter variansi. Penggunaannya menggunakan syarat bahwa distribusi data populasi adalah distribusi normal.

Teorema. Bila S^2 variansi sampel acak ukuran n diambil dari populasi normal dengan variansi σ^2, maka peubah acak

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

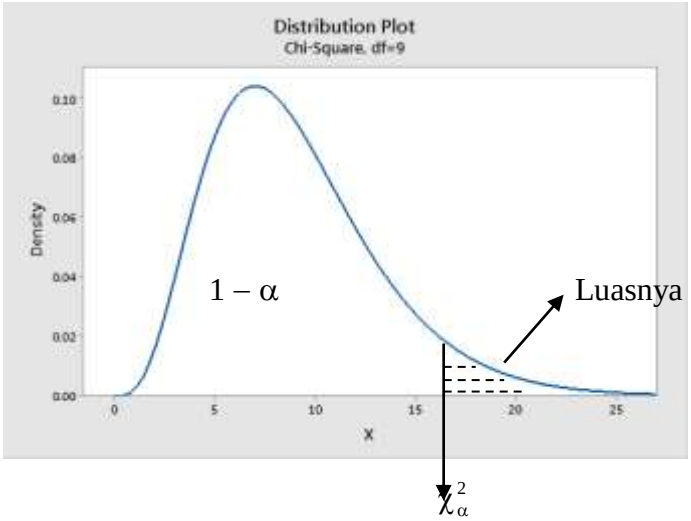
berdistribusi khi-kuadrat (*chi-square*) dengan derajat kebebasan $\nu = n - 1$ (ν dibaca "nu", dan χ^2 dibaca "khi-kuadrat"). Rataan dan variansi peubah acak ini ν dan 2ν. Pada Gambar 3.5.2.10 berikut diberikan contoh grafik Distribusi Khi-kuadrat dengan $\nu = 9$ dan $\nu = 14$.



Gambar 3.5.2.10 Grafik Distribusi Khi-kuadrat dengan $\nu = 9$ dan $\nu = 14$

Sama seperti halnya z_α pada distribusi normal, pada Distribusi Khi-kuadrat dikenal χ^2_α yang menyatakan nilai χ^2 di mana luas di sebelah kanannya adalah α seperti yang diperlihatkan pada Gambar 3.5.2.11 berikut. Penentuan nilai χ^2_α sama dengan seperti menentukan z_α pada distribusi normal, yaitu nilai χ^2 diperoleh dengan menggunakan peluang kumulatif sebagai berikut :

$$P(X^2 < \chi^2_\alpha) = 1 - \alpha$$

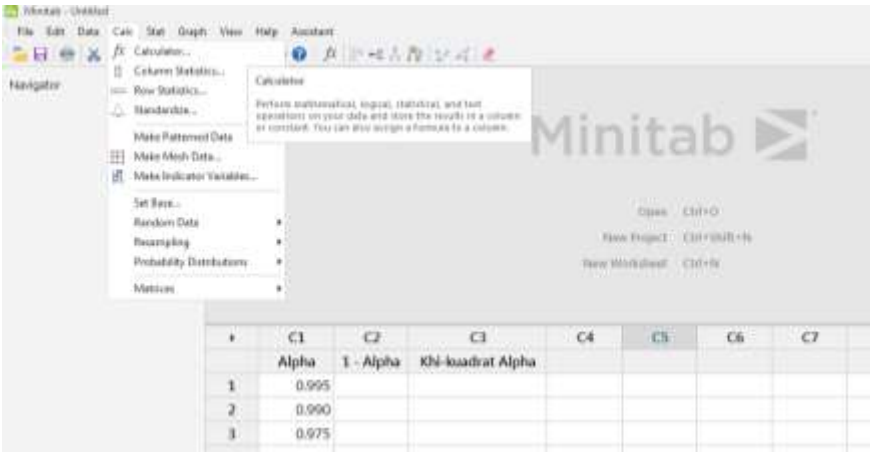


Gambar 3.5.2.11 Ilustrasi χ^2_α

Berikut ini diberikan langkah menentukan nilai χ^2_α untuk $\nu = 9$ dan beberapa α dengan menggunakan Program Minitab.

Langkah 1. Siapkan data pada kolom C1 Alpha dan isikan nilai Alpha (α) yang diinginkan.

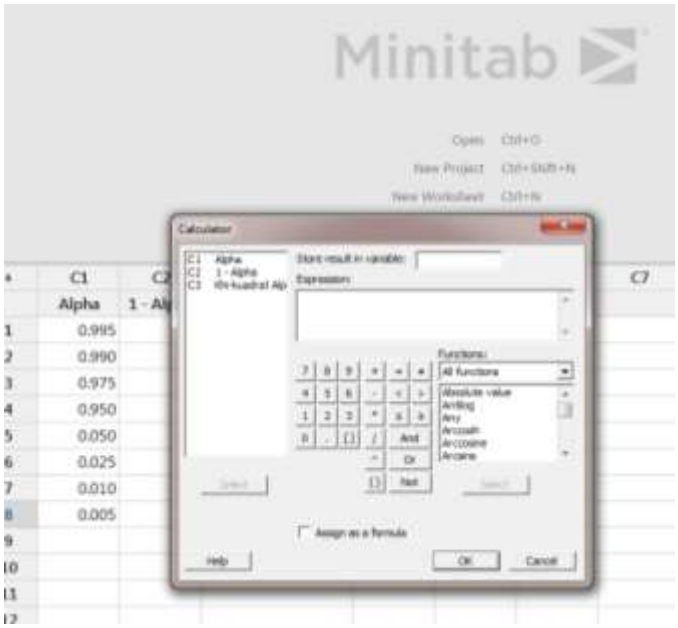
Siapkan kolom C2 1 – Alpha (Gambar 3.5.2.12)



Gambar 3.5.2.12 Input data α Normal

Langkah 2. Klik ikon *Calc*.

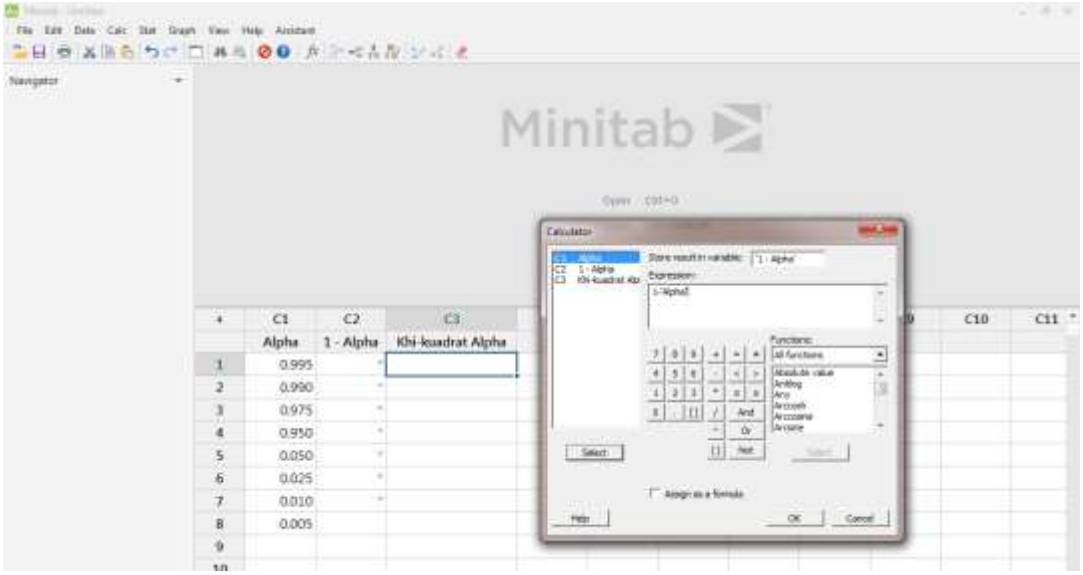
Maka akan muncul kotak *Calculator* seperti pada Gambar 3.5.13berikut.



Gambar 3.5.2.13 Kotak dialog *Calc*

Pada *Store result in variable* isikan C2 1 – Alpha. Lalu pada *Expression* ketikkan angka "1", kemudian lambang kurang yaitu "–", dan terkahir *Select* C1 Alpha dari kotak di sebelah kiri.

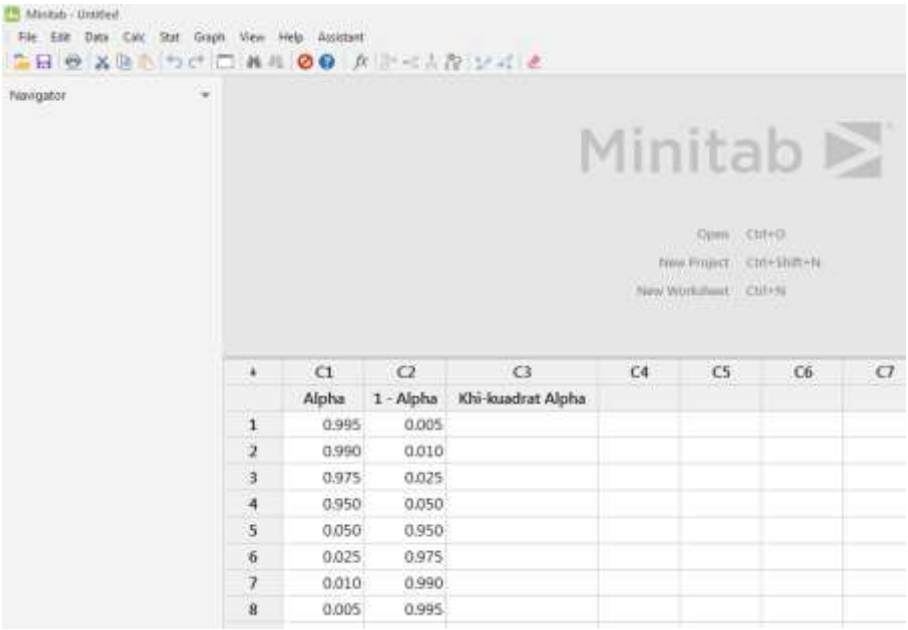
Expression ini bermakna bahwa kita sedang melakukan operasi pengurangan 1 – Alpha dan hasilnya akan disimpan pada C2 1 – Alpha (Gambar 3.5.2.14).



	C1	C2	C3
	Alpha	1 - Alpha	Khi-kuadrat Alpha
1	0.995		
2	0.990		
3	0.975		
4	0.950		
5	0.050		
6	0.025		
7	0.010		
8	0.005		
9			

Gambar 3.5.2.14 *Expression* untuk 1 – α

Kemudian klik Ok dan hasilnya seperti pada Gambar 3.5.2.15 berikut.



	C1	C2	C3	C4	C5	C6	C7
	Alpha	1 - Alpha	Khi-kuadrat Alpha				
1	0.995	0.005					
2	0.990	0.010					
3	0.975	0.025					
4	0.950	0.050					
5	0.050	0.950					
6	0.025	0.975					
7	0.010	0.990					
8	0.005	0.995					

Gambar 3.5.2.15 *Output Expression* 1 – α

Langkah 3. Klik ikon *Calc* dan *Probability Distributions*. Pilih *Chi-Square*. Pindahkan kursor ke *Inverse Cumulative Probability*. Isikan data-data seperti pada Gambar 3.5.2.16 berikut.

Gambar 3.5.2.16 *Inverse Cumulative Probability* Khi-kuadrat

Terakhir Ok. Maka akan tampil hasilnya seperti pada Gambar 3.5.2.17 berikut. Pada C3 Khi-kuadrat Alpha diberikan nilai χ^2_α.

	C1	C2	C3
	Alpha	1 - Alpha	Khi-kuadrat Alpha
1	0.995	0.005	1.7349
2	0.990	0.010	2.0879
3	0.975	0.025	2.7004
4	0.950	0.050	3.3251
5	0.050	0.950	16.9190
6	0.025	0.975	19.0228
7	0.010	0.990	21.6660
8	0.005	0.995	23.5894

Gambar 3.5.2.17 *Output Inverse Cumulative Probability* Khi-kuadrat

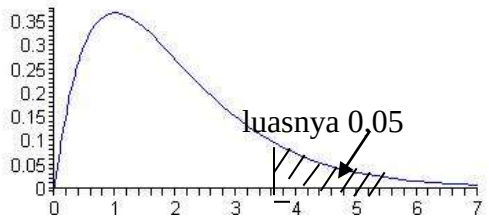
Nilai pada Gambar 3.5.2.17 di atas disusun ulang pada Pada Tabel 3.5.2.1 berikut.

Tabel 3.5.2.1 Tabel Khi-kuadrat dengan $\nu = 9$ untuk beberapa α

Alpha (α)	1 – Alpha ($1 - \alpha$)	Khi-kuadrat Alpha (χ^2_α)
0.995	0.005	1.734932905
0.99	0.01	2.087900736
0.975	0.025	2.7003895
0.95	0.05	3.325112843
0.05	0.95	16.9189776
0.025	0.975	19.0227678
0.01	0.99	21.66599433
0.005	0.995	23.58935078

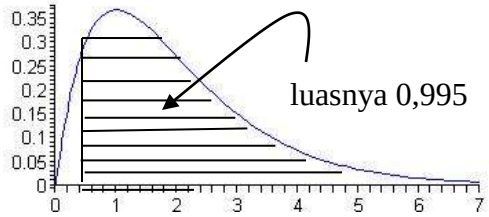
Dari tabel di atas, maka :

a. $\chi^2_{0,05,\nu=9} = 16{,}919$



Gambar 3.5.2.18 Nilai $\chi^2_{0,05,\nu=9}$

b. $\chi^2_{0,995,\nu=9} = 1{,}735$



Gambar 3.5.2.19 Nilai $\chi^2_{0,995,\nu=9}$

3.6. Distribusi Sampel Acak

Seperti telah dijelaskan pada bagian Pendahuluan bahwa dalam Statistika Inferensi kita akan melakukan inferensi (penentuan selang kepercayaan dan pengujian hipotesis) terhadap parameter populasi yang dilakukan berdasarkan keterangan dari sampel acak yang diambil dari populasi tersebut. Prosedur inferensi ini menggunakan konsep peluang, oleh karenanya dalam Ilmu Statistika perlu diketahui distribusi sampel acak yang diambil supaya kita dapat melakukan hitung peluang dalam prosedur inferensi tersebut.

Salah satu teorema yang sangat penting dalam konsep distribusi sampel acak adalah teorema kekonvergenan distribusi dan teorema yang paling umum digunakan adalah Teorema Limit Pusat (*Central Limit Theorem*). Berikut teorema tersebut.

Teorema. (Teorema Limit Pusat = TLP). Misalkan X adalah peubah acak dari suatu populasi dengan rata-rata μ dan variansi σ^2 yang berhingga. Misalkan pula $\overline{X}$ adalah rata-rata sampel acak ukuran n yang diambil dari populasi tersebut. Maka bila $n \to \infty$, limit distribusi dari peubah acak:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

adalah Normal Baku N(z; 0, 1). Limit distribusi Normal ini cukup baik sepanjang ukuran sampel $n \geq 30$, apapun bentuk distribusi populasi. Apabila ukuran sampel $n < 30$, maka harus diyakini distribusi populasi hampir normal. Dan jika populasi berdistribusi Normal, maka distribusi sampel acak $\overline{X}$ akan tepat berdistribusi Normal.

Pada pasal ini akan dibahas dua distribusi peubah acak kontinu yang secara teoritis diturunkan dari distribusi normal yaitu distribusi t dan distribusi F. Keduanya merupakan distribusi sampel acak seperti yang akan dibahas berikut ini.

a. Distribusi t

Dua parameter penting dalam suatu distribusi adalah rata-rata dan standar deviasi. Notasi rata-rata dan standar deviasi untuk populasi masing-masing adalah μ dan σ, sedangkan notasi rata-rata dan standar deviasi untuk sampel masing-masing adalah $\overline{X}$ dan s. Ketika kita mengambil sampel dari suatu populasi Normal dengan ukuran sampel n besar ($n \geq 30$), maka menurut Teorema Limit Pusat, distribusi rata-rata sampel akan mengikuti distribusi populasi yaitu berdistribusi Normal. Oleh karenanya dalam inferensi terhadap rata-rata sampel $\overline{X}$, standar deviasi sampel s dapat digunakan untuk menggantikan standar deviasi populasi σ, sehingga distribusi rata-rata sampel adalah,

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

Tetapi tidak demikian jika ukuran sampel n kecil ($n < 30$), maka distribusi rata-rata sampel tidak lagi berdistribusi Normal. Hal ini dikarenakan jika ukuran sampel kecil ($n < 30$), nilai s berubah cukup besar dari sampel ke sampel, sehingga sekarang kita tidak lagi menggunakan peubah acak menggunakan peubah acak $Z = (\overline{X} - \mu)/(\sigma/\sqrt{n})$ tetapi peubah acak $(\overline{X} - \mu)/(s/\sqrt{n})$ yang tidak lagi berdistribusi Normal. Distribusi $(\overline{X} - \mu)/(s/\sqrt{n})$ diturunkan sebagai berikut.

Pandang dua peubah acak kontinu yang saling bebas yaitu:

$$W = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

dan $V = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$ $\rightarrow$ Dari bentuk ini dapat kita tuliskan: $\frac{V}{n-1} = \frac{S^2}{\sigma^2}$

Sebut:

$$T = \frac{W}{\sqrt{V/(n-1)}} = \frac{(\overline{X}-\mu)/(\sigma/\sqrt{n})}{\sqrt{S^2/\sigma^2}} \text{ atau } T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$$

Yang bermakna bahwa distribusi $(\overline{X}-\mu)/(S/\sqrt{n})$ adalah distribusi T, yang distribusi peluangnya akan dikonsep dalam teorema berikut.

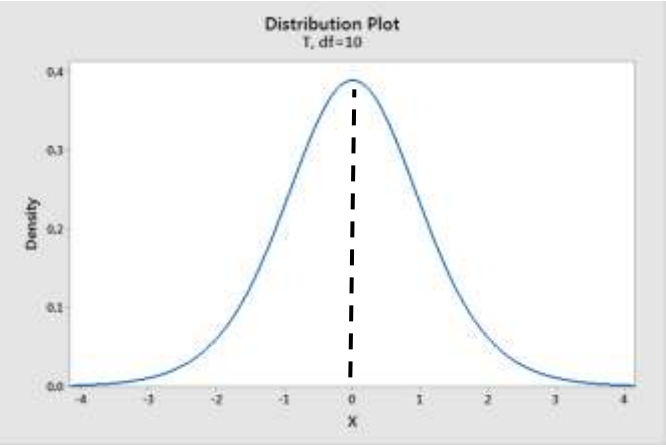
Teorema Misalkan W peubah acak normal baku dan V peubah acak khi-kuadrat dengan derajat kebebasan $\nu = n - 1$; seperti yang telah disebutkan di atas. Bila Z dan V bebas, maka distribusi peluang peubah acak T, bila

$$T = \frac{Z}{\sqrt{V/(n-1)}} = \frac{\overline{X}-\mu}{s/\sqrt{n}}$$

diberikan oleh

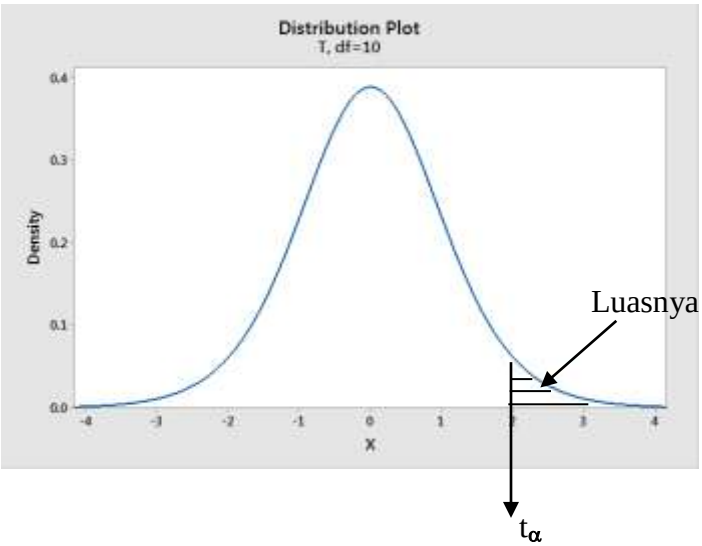
$$h(t) = \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2} \qquad -\infty < t < \infty$$

dengan derajat kebebasan (*degrees of freedom* = df) $\nu = n - 1$. Kurva Distribusi T mirip dengan kurva Distribusi Normal yang berbentuk lonceng dan simetris terhadap rata-ratanya yang sama-sama bernilai 0. Pada Gambar 3.6.1 berikut diberikan grafik Distribusi T dengan df = 10.



Gambar 3.6.1 Grafik T dengan df = 10

Sama halnya seperti Distribusi Normal dan Distribusi Khi-kuadrat, dalam Statistika Inferensi kita akan menggunakan nilai t_α yaitu nilai peubah acak T sehingga luas di kanannya adalah α, seperti yang diperlihatkan pada Gambar 3.6.2 berikut.



Gambar 3.6.2 Ilustrasi t_α

Dengan cara yang sama seperti penjelasan pada Distribusi Khi-kuadrat (halaman), kita dapat menentukan nilai t_α untuk derajat bebas ν tertentu dan α tertentu (tidak dijelaskan lagi). Pada (Gambar 3.6.3) diberikan nilai $t_{\alpha,\,\nu = 10}$ (C3) untuk beberapa α (C1).

	C1	C2	C3	C4
	Alpha	1 - Alpha	t - Alpha	
1	0.050	0.950	1.81246	
2	0.025	0.975	2.22814	
3	0.010	0.990	2.76377	
4	0.005	0.995	3.16927	
5				
6				

Gambar 3.6.3 Nilai $t_{\alpha,\,\nu = 10}$ untuk beberapa α

b. Distribusi F

Distribusi sampel acak yang kedua adalah Distribusi F. Pada dasarnya distribusi ini digunakan ketika kita melakukan prosedur Statistika Inferensi dalam membandingkan dua variansi populasi. Dalam aplikasi yang lebih luas, distribusi F juga digunakan pada prosedur pengujian dalam Analisis Regresi. Konsep Distribusi F diberikan pada teorema berikut.

Teorema. Misalkan U dan V dua peubah acak bebas masing-masing berdistribusi khi-kuadrat dengan derajat kebebasan ν_1 dan ν_2. Maka distribusi peluang peubah acak F yaitu :

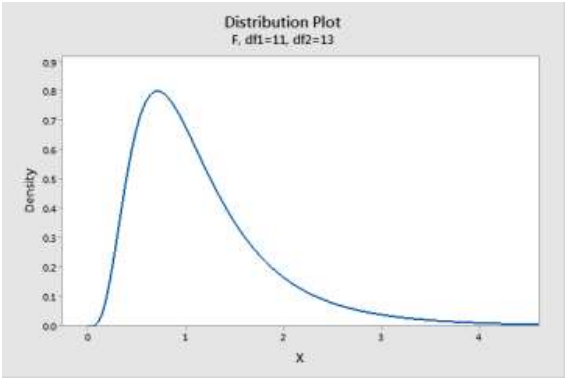
$$F = \frac{U/\nu_1}{V/\nu_2}$$

diberikan oleh

$$h(f) = \frac{\Gamma[(\nu_1+\nu_2)/2](\nu_1/\nu_2)^{\nu_1/2}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \quad \frac{f^{(\nu_1/2)-1}}{[1+(\nu_1/\nu_2)f]^{(\nu_1+\nu_2)/2}} \quad , \ 0 < f < \infty$$

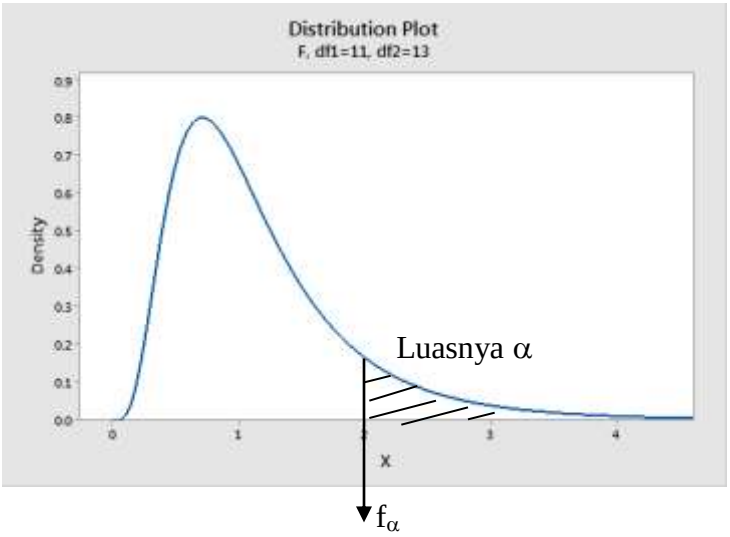
$$= \qquad 0 \qquad , \text{ untuk f lainnya}$$

dengan derajat kebebasan ν_1 dan ν_2. Contoh grafik f dengan ν_1 (df1) = 11 dan ν_2 (df2) = 13, diberikan pada Gambar 3.6.4 berikut.



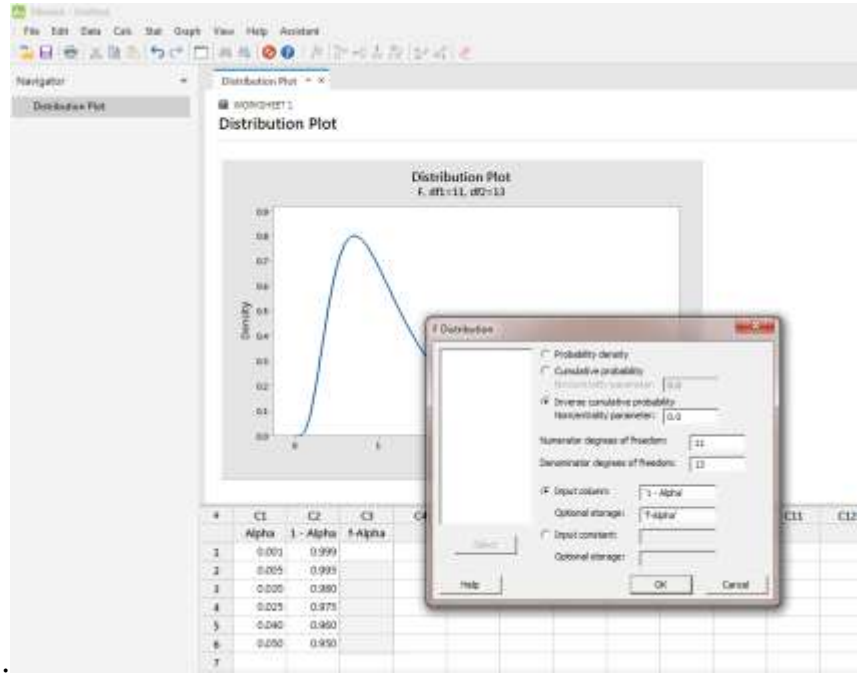
Gambar 3.6.4 Grafik F dengan $\nu_1 = 11$, $\nu_2 = 13$

Sama halnya seperti tiga distribusi peubah acak kontinu yang telah disebutkan sebelumnya yaitu Distribusi Normal, Distribusi Khi-kuadrat dan Distribusi T, dalam penggunaannya kita harus menentukan nilai f_α yaitu nilai peubah acak F sehingga di sebelah kanannya terdapat luas sebesar α (Gambar 3.6.5). Untuk menentukan nilai f_α untuk berbagai ν_1, ν_2 dan α dapat dilakukan dengan menggunakan Program Minitab.



Gambar 3.6.5. Ilustrasi f_α

Berikut ini diberikan gambar nilai f_α untuk ν_1 (df1) = 11 dan ν_2 (df2) = 13, dan beberapa α. Langkahnya sama seperti yang dilakukan pada Distribusi Khi-kuadrat (Gambar 3.6.6 dan Gambar 3.6.7).



Gambar 3.6.6 Langkah awal menentukan f_α dengan Program Minitab

	C1	C2	C3
	Alpha	1 - Alpha	f-Alpha
1	0.001	0.999	6.64738
2	0.005	0.995	4.72405
3	0.020	0.980	3.38958
4	0.025	0.975	3.19750
5	0.040	0.960	2.81059
6	0.050	0.950	2.63465
7			

Gambar 3.6.7 Nilai $f_{\alpha,\ v1\ =\ 11,\ v2\ =\ 13}$ untuk beberapa α

4. Penaksiran

4.1. Pengantar.

Sekarang kita akan membahas konsep yang mendasari penggunaan Ilmu Statistika dalam pengolahan data yaitu konsep Statistika Inferensi. Dalam Statistika Inferensi yang menjadi perhatian bagaimana keterangan (lebih lanjut disebut karakteristik) pada sampel dapat digunakan untuk memberikan kesimpulan tentang karakteristik populasi. Seperti telah dijelaskan pada pasal Pendahuluan, karakteristik sampel disebut statistik dan karakteristik populasi disebut parameter. Dalam pasal ini akan dibahas bagaimana sebuah statistik sampel dapat digunakan sebagai penduga atau penaksir bagi suatu parameter populasi, yang kesemuanya tentu saja didasarkan pada konsep peluang. Dalam konteks ini perlu dipahami hal-hal berikut ini:

1. Populasi dinyatakan sebagai peubah acak X yang mempunyai distribusi peluang tertentu dengan rata-rata μ dan variansi σ^2.
2. Populasi mempunyai satu atau lebih parameter yang tidak diketahui, diberi notasi θ.
   Secara umum notasi parameter populasi dituliskan dalam huruf Yunani (Greek).
3. Fungsi peluang peubah acak X (populasi) dituliskan sebagai berikut:
   a. $f(x; \boldsymbol{\theta})$ untuk peubah acak kontinu
   b. $p(x; \boldsymbol{\theta})$ untuk peubah acak diskrit

   Penulisan $\boldsymbol{\theta}$ di sini tebal sebagai lambang bahwa parameter θ mungkin lebih dari satu, yaitu:

   $$\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_p] \in \Omega \subset R^p \text{ , } p \geq 1$$

   Di mana : $\theta_1, \theta_2, \ldots, \theta_p$ semuanya tidak diketahui

   Dan R^p adalah ruang vektor real berdimensi-p

   Jika parameter θ hanya satu, kita dapat menuliskannya tidak tebal (seperti contoh berikut).

   Contoh.

   1. Distribusi Binomial : $X \sim b(x; n, p)$, p tidak diketahui
      Di sini: $\theta = p$ yaitu proporsi sukses dalam populasi.
   2. Distribusi Normal : $X \sim N(x; \mu, \sigma^2)$, μ tidak diketahui, σ^2 diketahui
      Di sini: $\theta = \mu$ rata-rata populasi
   3. Distribusi Normal : $X \sim N(x; \mu, \sigma^2)$, μ dan σ^2 keduanya tidak diketahui
      Di sini: $\boldsymbol{\theta} = (\mu, \sigma^2) \in R^2$ yaitu: $\theta_1 = \mu$ rata-rata populasi dan $\theta_2 = \sigma^2$ variansi populasi

4. Untuk menaksir parameter θ digunakan statistik sampel acak yang diambil dari populasi.
5. Sampel acak ukuran n dituliskan $X_1, X_2, \ldots, X_n$, di mana:
   a. Distribusi peluang sampel acak $X_1, X_2, \ldots, X_n$ identik/sama dengan distribusi populasinya, artinya rata-rata dan variansi masing-masing sampel acak $X_1, X_2, \ldots, X_n$ sama dengan mean dan variansi populasinya.
   b. $X_1, X_2, \ldots, X_n$ adalah sampel acak yang saling bebas.
6. Pemahaman sampel acak seperti yang disebutkan pada bagian nomor 5 disebut sampel acak yang *independent and identically distributed* (iid).

Tinjau peubah acak X dengan fungsi peluang f(x; θ) atau p(x; θ). Misalkan $T = T(X_1, X_2, \ldots, X_n)$ merupakan fungsi dari sampel acak $X_1, X_2, \ldots, X_n$. Fungsi $T = T(X_1, X_2, \ldots, X_n)$ akan digunakan untuk menaksir θ, yang disebut penaksir titik bagi θ. Selanjutnya fungsi $T = T(X_1, X_2, \ldots, X_n)$ yang demikian disebut statistik. Jadi statistik adalah fungsi dari sampel acak yang digunakan untuk menaksir parameter θ.

Contoh.

1. Untuk sampel acak $X_1, X_2, ..., X_n$ yang berasal populasi dengan rata-rata μ dan variansi σ^2 ; maka penaksir titik untuk μ dan σ^2 (penaksir biasanya ditulis dengan huruf Greek dengan topi di atasnya atau huruf Latin yang sesuai) :

$$\hat{\mu} = T(X_1, X_2, ..., X_n) = \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{1}{n}\left(X_1 + X_2 + ... + X_n\right)$$

$$\hat{\sigma}^2 = T(X_1, X_2, ..., X_n) = s^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1} = \frac{1}{n-1}\left[(X_1 - \overline{X})^2 + ... + (X_n - \overline{X})^2\right]$$

   yang merupakan rata-rata dan variansi sampel acak iid $X_1, X_2, ..., X_n$.

2. Dalam percobaan Binomial dengan jumlah ulangan n dan peluang sukses adalah p. Maka penaksir untuk p adalah :

$$\hat{p} = T(X_1, X_2, ..., X_n) = \frac{\sum_{i=1}^{n} X_i}{n} \text{ ,, dengan } \sum_{i=1}^{n} X_i \text{ banyaknya sukses terjadi}$$

4.2. Penaksir Kemungkinan Maksimum

Dalam pasal ini kita akan membahas salah satu metode yang digunakan untuk menaksir dua parameter terpenting dalam pengolahan data, yaitu rata-rata μ dan variansi σ^2. Metode ini disebut metode kemungkinan maksimum (*maximum likelihood method*) yang bermakna penentuan penaksir parameter dilakukan dengan prinsip bahwa kemungkinan atau peluang untuk memperoleh penaksir tersebut adalah maksimum. Dalam konsep yang akan dijelaskan berikut ini, kita menggunakan fungsi peluang peubah acak kontinu $f(x; \theta)$. Konsep tersebut akan berlaku juga untuk peubah acak diskrit dengan mengganti $f(x; \theta)$ menjadi $p(x; \theta)$.

Tinjau peubah acak X dengan fungsi peluang $f(x; \theta)$, $\theta \in \Omega$, rataan μ dan variansi σ^2. Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak (iid) dari peubah acak X. Oleh karena distribusi sampel acak ini identik dengan peubah acak X, maka fungsi peluang dari masing-masing sampel acak adalah $f(x_i; \theta)$. Kemudian kita nyatakan fungsi peluang yang dinamakan fungsi kemungkinan (*likelihood function*) sebagai berikut

$$L(\theta, \mathbf{x}) = \prod_{i=1}^{n} f(x_i, \theta) = f(x_1, \theta).f(x_2, \theta). \ldots . f(x_n, \theta), \theta \in \Omega$$

di mana $\mathbf{x} = [x_1 \; x_2 \; \ldots \; x_n]^T$. Selanjutnya dalam rangka menentukan penaksir untuk θ, fungsi kemungkinan $L(\theta, \mathbf{x})$ disingkat dengan $L(\theta)$. Sebut fungsi $\ell(\theta) = \log L(\theta)$ yaitu

$$\ell(\theta) = \log L(\theta) = \log[f(x_1, \theta).f(x_2, \theta). \ldots . f(x_n, \theta)] = \sum_{i=1}^{n} \log f(x_i, \theta)$$

Metode penaksiran parameter θ dilakukan secara Kalkulus yaitu dengan menggunakan fungsi turunan pertama untuk menentukan nilai kritis θ dan menggunakan fungsi turunan kedua untuk menunjukkan bahwa $L(\theta)$ mencapai maksimum pada nilai kritis θ yang diperoleh. Persamaan untuk memperoleh nilai kritis θ, selanjutnya disebut persamaan estimasi (*estimation equation*) diberikan oleh

$$\frac{d\,\ell(\theta, \mathbf{x})}{d\theta} = 0 \qquad (4.2.1)$$

Penaksir yang diperoleh dari Persamaan (4.2.1) di atas ditulis $\hat{\theta}_{mle}$ yaitu penaksir kemungkinan maksimum (*maximum likelihood estimator*).

Contoh 1. (Kasus 1 parameter) Tinjau peubah acak diskrit X dengan distribusi peluang disebut Distribusi Bernoulli dengan fungsi massa peluang sebagai berikut.

$$p(x,\theta)=\begin{cases}\theta^{x}(1-\theta)^{1-x} & x=0,1\\ 0 & x\ \text{lain}\end{cases};\ 0\le\theta\le 1$$

Maka:

$$L(\theta,x)=\prod_{i=1}^{n}p(x_i,\theta)=\prod_{i=1}^{n}\theta^{x_i}(1-\theta)^{1-x_i}\quad=\theta^{\sum_1^n x_i}(1-\theta)^{n-\sum_1^n x_i}$$

Dan :

$$\ell(\theta)=\ln L(\theta)=\left(\sum_1^n x_i\right)\ln\theta+\left(n-\sum_1^n x_i\right)\ln(1-\theta)$$

Maka persamaan estimasi:

$$\frac{d\ell(\theta)}{d\theta}=\frac{\sum_1^n x_i}{\theta}-\frac{n-\sum_1^n x_i}{1-\theta}=0$$

Kiri kanan kali $\theta(1-\theta)$

$$(1-\theta)\sum_1^n x_i-\theta\left(n-\sum_1^n x_i\right)=0$$

$$\sum_1^n x_i-n\theta=0\rightarrow n\theta=\sum_1^n x_i\rightarrow\hat{\theta}_{mle}=\frac{\sum_1^n x_i}{n}=\overline{X}$$

Untuk menunjukkan bahwa $L(\theta)$ mencapai maksimum pada $\hat{\theta}_{mle}=\overline{X}$ kita dapat memeriksa bahwa :

$$\frac{d^2\ell(\theta)}{d\theta^2}<0$$

Dari : $\frac{d\ell(\theta)}{d\theta}=\frac{\sum_1^n x_i}{\theta}-\frac{n-\sum_1^n x_i}{1-\theta}$ diperoleh : $\frac{d^2\ell(\theta)}{d\theta^2}=-\frac{\sum_1^n x_i}{\theta^2}-\frac{n-\sum_1^n x_i}{(1-\theta)^2}$

Kemudian subsitusikan $\hat{\theta}_{mle}=\overline{X}$ dalam θ pada persamaan di atas.

Jika $\hat{\theta}_{mle} = \frac{\sum_1^n x_i}{n} = \overline{X}$ disubsitusikan ke dalam persamaan $-\frac{\sum_1^n x_i}{\theta^2} - \frac{n - \sum_1^n x_i}{(1-\theta)^2}$, diperoleh :

$$-\frac{n}{\sum_1^n x_i} - \frac{n - \sum_1^n x_i}{\left(1 - \frac{\sum_1^n x_i}{n}\right)^2} < 0$$

$\longrightarrow$ Penyebut bentuk kuadrat, pasti > 0

Sebab dalam Distribusi Bernoulli nilai x adalah 0 atau 1, sehingga $0 \le \sum_1^n x_i \le n$ sehingga $n - \sum_1^n x_i \ge 0$. Oleh karena itu $\frac{d^2 l(\theta)}{d\theta^2} < 0$ pada $\hat{\theta}_{mle} = \overline{X}$.

Contoh 2. (Kasus 2 parameter). Tinjau peubah acak kontinu dengan distribusi peluang adalah Distribusi Normal X ~ N(x; θ_1, θ_2), dimana θ_1 rata-rata populasi ($-\infty < \theta_1 < \infty$) dan θ_2 variansi populasi ($\theta_2 > 0$) yang keduanya tidak diketahui dan akan ditaksir dengan menggunakan metode kemungkinan maksimum. Fungsi peluang distribusi normal adalah:

$$n(x;\theta_1,\theta_2) = \frac{1}{\sqrt{2\pi\theta_2}}.\exp\left[-\frac{1}{2\theta_2}(x=\mu)^2\right], -\infty < x < \infty$$

Ambil sampel acak iid $X_1, X_2, \ldots, X_n$, maka:

$$L(\theta_1,\theta_2,x) = \prod_{i=1}^n n(x_i,\theta_1,\theta_2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta_2}}.\exp\left[-\frac{1}{2\theta_2}(x_i=\mu)^2\right]$$

$$= \left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n.\exp\sum_{i=1}^n\left[-\frac{1}{2}\left(\frac{x_i-\theta_1}{\theta_2}\right)^2\right]$$

Dan $\ell(\theta)$:

$$\ell(\theta_1,\theta_2) = -\frac{n}{2}\ln\theta_2 - \frac{n}{2}\ln(2\pi) - \frac{1}{2\theta_2}\sum_{i=1}^n (x_i-\theta_1)^2$$

Dalam Distribusi Normal ada 2 parameter yang akan ditaksir, maka kita akan menggunakan Kalkulus Fungsi Dua Variabel dalam menentukan kedua penaksir tersebut. Persamaan penaksir adalah

$$\frac{d\ell(\theta_1,\theta_2)}{d\theta_1}=0 \text{ dan } \frac{d\ell(\theta_1,\theta_2)}{d\theta_2}=0$$

Penaksir θ_1.

$$\frac{d\ell(\theta_1,\theta_2)}{d\theta_1}=\frac{d}{d\theta_1}\left[-\frac{n}{2}\ln\theta_2-\frac{n}{2}\ln(2\pi)\ -\frac{1}{2\theta_2}\sum_{i=1}^{n}(x_i-\theta_1)^2\right]=0$$

$$-\frac{1}{2\theta_2}.(2).(-1).\sum_{i=1}^{n}(x_i-\theta_1)=0\rightarrow\sum_{i=1}^{n}(x_i-\theta_1)=\sum_{i=1}^{n}x_i-n\theta_1=0\rightarrow\hat{\theta}_{1,mle}=\frac{\sum_{1}^{n}x_i}{n}=\overline{X}$$

Penaksir θ_2.

$$\frac{d\ell(\theta_1,\theta_2)}{d\theta_2}=\frac{d}{d\theta_2}\left[\ n\ln\left(\frac{1}{\theta_2\sqrt{2\pi}}\right)-\frac{1}{2\theta_2}\sum_{i=1}^{n}(x_i-\theta_1)^2\right]=0$$

$$\frac{d}{d\theta_2}\left[-\frac{n}{2}\ln\theta_2-\frac{n}{2}\ln(2\pi)\ -\frac{1}{2\theta_2}\sum_{i=1}^{n}(x_i-\theta_1)^2\right]=0$$

$$-\frac{n}{2\theta_2}+\frac{1}{2\theta_2^2}\sum_{i=1}^{n}(x_i-\theta_1)^2=0\rightarrow-n\theta_2+\sum_{i=1}^{n}(x_i-\theta_1)^2=0\rightarrow\widehat{\theta}_2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\theta_1)^2$$

Dengan mensubsitusikan

$$\hat{\theta}_{1,mle}=\frac{\sum_{1}^{n}x_i}{n}=\overline{X}$$

Maka penaksir kemungkinan maksimum untuk θ_2

$$\widehat{\theta}_{2,mle}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$$

Dapat dibuktikan bahwa syarat turunan kedua untuk kedua nilai kritis ini dipenuhi, bahwa:

$$\frac{d^2\ell(\theta_1,\theta_2)}{d\theta_1^2}<0 \text{ dan } \frac{d^2\ell(\theta_1,\theta_2)}{d\theta_2^2}<0$$

Yaitu: $\dfrac{d^2\ell(\theta_1,\theta_2)}{d\theta_1^2} = \dfrac{1}{\theta_2}\sum_{i=1}^{n}(x_i-\theta_1) = -1 < 0$

Akan selalu negatif untuk setiap θ_1 termasuk pada : $\hat{\theta}_{1,mle} = \dfrac{\sum_{1}^{n} x_i}{n} = \overline{X}$

Dan:

$$\frac{d^2\ell(\theta_1,\theta_2)}{d\theta_2^2} = \frac{d}{d\theta_2}\left[-\frac{n}{2\theta_2}+\frac{1}{2\theta_2^2}\sum_{i=1}^{n}(x_i-\theta_1)^2\right] = -\frac{n}{2\theta_2^2}-\frac{1}{2\theta_2^3}\sum_{i=1}^{n}(x_i-\theta_1)^2 < 0$$

Akan selalu bernilai negatif sebab θ_2 adalah σ^2 yang nilainya selalu positif termasuk pada

$$\widehat{\theta}_{2,mle} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$$

Syarat lain yang harus diperiksa adalah determinan:

$$\begin{vmatrix} \dfrac{d^2\ell(\theta_1,\theta_2)}{d\theta_1^2} & \dfrac{d^2\ell(\theta_1,\theta_2)}{d\theta_1\,d\theta_2} \\ \dfrac{d^2\ell(\theta_1,\theta_2)}{d\theta_1\,d\theta_2} & \dfrac{d^2\ell(\theta_1,\theta_2)}{d\theta_2^2} \end{vmatrix}$$ harus positif

Dari : $\dfrac{d\ell(\theta_1,\theta_2)}{d\theta_1} = \sum_{i=1}^{n}(x_i-\theta_1) \rightarrow \dfrac{d^2\ell(\theta_1,\theta_2)}{d\theta_1\,d\theta_2} = 0$

Maka:

$$\begin{vmatrix} -1 & 0 \\ 0 & -\dfrac{n}{\theta_2^2}-\dfrac{1}{\theta_2^3}\sum_{i=1}^{n}(x_i-\theta_1)^2 \end{vmatrix} = \frac{n}{\theta_2^2}+\frac{1}{\theta_2^3}\sum_{i=1}^{n}(x_i-\theta_1)^2 > 0$$

Demikianlah telah diberikan dua contoh cara menentukan penaksir untuk suatu parameter dengan menggunakan metode kemungkinan maksimum. Selanjutya akan dibahas kriteria penaksir untuk suatu parameter.

4.3. Kriteria Penaksir

Sekarang kita akan membahas konsep dasar yang menjadi kriteria suatu penaksir bagi parameter populasi yang diberikan dalam definisi-definisi berikut. Kriteria ini akan menyangkut masalah konsep rata-rata (ekspektasi) dan variansi dari penaksir tersebut.

Definisi 1. Tinjau peubah acak X dengan fungsi peluang f(x; θ) atau p(x; θ), θ ∈ Ω. Misalkan statistik T = T(X_1, X_2, …, X_n) merupakan fungsi dari sampel acak X_1, X_2, …, X_n. Maka Fungsi T = T(X_1, X_2, …, X_n) disebut penaksir tak bias untuk θ jika

E[T] = θ, untuk setiap θ ∈ Ω

Jika E[T] ≠ θ maka T disebut penaksir bias untuk θ.

Pada halaman 22 telah dijelaskan bahwa konsep ekspektasi adalah rata-rata. Dengan demikian T disebut penaksir tak bias apabila nilai rata-rata T sama dengan nilai parameter θ yang akan ditaksir. Hal ini merupakan logika yang mendasar dari suatu penaksir yang tak bias.

Contoh.

1. Misalkan X adalah peubah acak dengan fungsi peluang f(x; θ) atau p(x; θ), θ ∈ Ω, dengan rata-rata θ dan variansi σ^2. Misalkan juga X_1, X_2, …, X_n adalah sampel acak iid dari peubah acak X, maka masing-masing sampel acak mempunyai rata-rata θ dan variansi σ^2. Kita akan memperlihatkan bahwa penaksir kemungkinan maksimum $\hat{\theta}_{mle} = \overline{X}$ adalah penaksir tak bias untuk rata-rata populasi θ, yang dijelaskan sebagai berikut.

$$E\left[\overline{X}\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\theta = \frac{1}{n}.n\theta = \theta$$

Jadi $\overline{X}$ adalah penaksir tak bias untuk θ, dengan variansi sebagai berikut:

$$\text{var}\left(\overline{X}\right) = \text{var}\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n}\text{var}\left[X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}.n\sigma^2 = \frac{\sigma^2}{n}$$

2. Dalam Contoh 2 ini akan kita perlihatkan bahwa penaksir kemungkinan maksimum untuk variansi dalam Contoh 2 halaman 59 merupakan penaksir bias dan kita akan menentukan penaksir tak bias untuk variansi.

$$E\left[\hat{\theta}_{2,mle}\right] = E\left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}\right]$$

Untuk menentukan ekspektasi ini kita lihat penjelasan berikut.

Pada halaman 43 telah dijelaskan bahwa hal-hal yang berkaitan dengan variansi akan menggunakan Distribusi Khi-kuadrat, yang dituliskan sebagai berikut:

$$X^2 = \frac{(n-1)\,S^2}{\sigma^2} \quad \text{Atau :} \ \frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}$$

Pada halaman 44 disebutkan bahwa rata-rata atau ekspektasi $X^2 = \frac{n-1}{\sigma^2}S^2$ adalah $\nu = n - 1$ yaitu:

$$E[X^2] = E\left[\frac{(n-1)\,S^2}{\sigma^2}\right] = n-1$$

Karena E dalam persamaan di atas adalah operator linear, maka dari sini diperoleh:

$\frac{n-1}{\sigma^2}E\left[S^2\right] = n-1 \rightarrow E[S^2] = \sigma^2$ yang menyatakan bahwa S^2 penaksir tak bias untuk σ^2.

Di mana : $S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$

Bagaimana dengan $\hat{\theta}_{2,mle}$??

$$\hat{\theta}_{2,mle} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n} = \frac{n-1}{n}\,\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} = \frac{n-1}{n}S^2$$

Maka :

$$E\left[\hat{\theta}_{2,mle}\right] = E\left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}\right] = E\left[\frac{n-1}{n}\,\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}\right] = E\left[\frac{n-1}{n}S^2\right] = \frac{n-1}{n}E[S^2] = \frac{n-1}{n}\sigma^2$$

Dari sini kita tahu bahwa $\widehat{\theta}_{2,\text{mle}}$ adalah penaksir bias untuk σ^2. Itu sebabnya dalam aplikasi, rumus variansi yang digunakan adalah S^2 yang merupakan penaksir tak bias untuk σ^2.

Statistik T yang merupakan fungsi dari sampel acak, merupakan suatu fungsi yang tergantung pada ukuran sampel n. Dalam konteks yang akan dibahas berikut ini, fungsi T dituliskan T_n di mana n adalah ukuran sampel yang nilainya semakin besar, yang ditulis dengan notasi $n \to \infty$, sehingga sekarang kita mempunyai apa yang dinamakan barisan statistik T_n. Yang menjadi perhatian adalah bagaimana distribusi peluang dari statistik T_n (yang merupakan barisan peubah acak menurut $n \to \infty$) ketika ukuran sampel n semakin besar. Dari sinilah muncul pengertian kekonvergenan dalam peluang.

Definisi. Misalkan $\{X_n\}$ adalah barisan peubah acak dan X adalah peubah acak yang terdefinisi pada ruang sampelnya. Barisan $\{X_n\}$ dikatakan konvergen dalam peluang ke X jika untuk setiap $\varepsilon > 0$,

$$\lim_{n \to \infty} P[\,|X_n - X| \geq \varepsilon\,] = 0$$

atau ekivalen dengan,

$$\lim_{n \to \infty} P[\,|X_n - X| < \varepsilon\,] = 1$$

Ditulis dengan notasi,

$$X_n \xrightarrow{P} X$$

Teorema (*Weak Law of Large Numbers* = WLLN). Misalkan $\{X_n\}$ adalah barisan peubah acak iid dengan rata-rata yang sama μ dan variansi $\sigma^2 > 0$. Misalkan,

$$\overline{X}_n = \frac{\sum_{1}^{n} X_i}{n}$$

maka : $\overline{X}_n \xrightarrow{P} \mu$.

Jadi selain $\overline{X} = \frac{\sum_{1}^{n} x_i}{n}$ merupakan penaksir tak bias untuk μ, juga dengan Teorema WLLN ini diperlihatkan bahwa jika $n \to \infty$, maka barisan peubah acak $\overline{X}_n = \frac{\sum_{1}^{n} x_i}{n}$ akan konvergen ke μ. Hal ini bermakna semakin besar nilai n, maka secara peluang, semakin dekat nilai $\overline{X}_n$ ke μ.

Teorema (Chebyshev's *Inequality*). Misalkan X adalah suatu peubah acak dengan rataan μ dan variansi $\sigma^2 < \infty$. Maka untuk setiap bilangan $k > 0$,

$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$ atau ekivalen dengan: $P[|X - \mu| < k\sigma] \geq 1 - \frac{1}{k^2}$

Definisi kekonvergenan dalam peluang dan teorema Chebyshev's *Inequality* merupakan konsep yang mendasari kriteria kedua suatu penaksir parameter. Kedua teorema menggambarkan bahwa jika n sangat besar, maka nilai penaksir tak bias akan sangat mendekati nilai parameter yang sesungguhnya. Hal ini nantinya mengarah kepada konsep variansi penaksir tak bias tersebut, yang diharapkan nilainya menuju nol ketika n menuju tak hingga.

Contoh. Kita tinjau penaksir kemungkinan maksimum untuk rata-rata μ.

$$\overline{X}_n = \frac{\sum_{1}^{n} x_i}{n}, \text{ dengan } \operatorname{var}(\overline{X}) = \frac{\sigma^2}{n}$$

Menurut Definisi Kekonvergenan dalam peluang dan Teorema Chebyshev's *Inequality*

$$P\left(\left|\overline{X} - \mu\right| \geq \varepsilon . \frac{\sigma/\sqrt{n}}{\sigma/\sqrt{n}}\right) = P\left(\left|\overline{X} - \mu\right| \geq \varepsilon . \underbrace{\frac{\sqrt{n}}{\sigma}}_{= k} . \underbrace{\sigma/\sqrt{n}}_{= \sigma_{\overline{X}}}\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

Maka :

$$\lim_{n \to \infty} P\left(\left|\overline{X} - \mu\right| \geq \varepsilon . \frac{\sqrt{n}}{\sigma} . \sigma/\sqrt{n}\right) = \lim_{n \to \infty} \frac{\sigma^2}{n\varepsilon^2} = 0$$

Hal ini merupakan cara untuk memeriksa apakah suatu barisan peubah acak konvergen dalam peluang terhadap suatu peubah acak tertentu, yaitu dengan melihat limit variansinya apakah menuju nol ketika n menuju tak hingga, seperti yang diperlihatkan pada contoh di atas. Hal ini akan kita nyatakan sebagai konsep konsistensi yang dinyatakan dalam definisi berikut.

Definisi. Misalkan X adalah suatu variabel random dengan fungsi distribusi kumulatif $(x; \theta)$, $\theta \in \Omega$. Misalkan juga $X_1, X_2, \ldots, X_n$ adalah sampel acak iid dan T_n adalah statistik. Statistik T_n disebut sebagai penaksir konsisten untuk parameter θ jika

$$T_n \xrightarrow{P} \theta$$

Berdasarkan definisi ini dan pembahasan di atas, dan juga berdasarkan teorema WLLN, maka kita katakan bahwa $\overline{X}_n$ penaksir konsisten untuk μ.

Contoh. Akan ditunjukkan bahwa S^2 adalah penaksir konsisten untuk σ^2.
Dari:

$$\text{var}\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1)$$

$$\frac{(n-1)^2}{\sigma^4}\text{var}\,(S^2) = 2(n-1)$$

$$\text{var}\,(S^2) = \frac{2(n-1)}{(n-1)^2}.\sigma^4 = \frac{2\sigma^4}{n-1} \rightarrow 0$$

4.4. Penaksiran Parameter

Aktivitas dalam Statistika Inferensi terdiri dari dua yaitu penaksiran parameter populasi dan pengujian hipotesis terhadap parameter populasi. Kita mulai dengan konsep penaksiran parameter populasi.

Dalam konsep penaksiran parameter populasi, kita mulai dengan dua parameter yang menggambarkan ukuran pemusatan data yaitu rataan μ, dan yang menggambarkan ukuran penyebaran data yaitu variansi σ^2. Pada Tabel 4.4.1 berikut diberikan empat karakteristik data, notasi dalam populasi (parameter) dan notasi dalam sampel (statistik).

Tabel 4.4.1 Karakteristik Data

Karakteristik	PARAMETER Karakteristik Populasi (teoritis)	STATISTIK Karakteristik Sampel (empiris)
Rataan	μ	$\hat{\mu} = \overline{X}$
Variansi	σ^2	$\hat{\sigma}^2 = s^2$
Standar deviasi	σ	$\hat{\sigma} = s$
Proporsi	p	$\hat{p}$

Seperti telah dijelaskan sebelumnya, statistik adalah fungsi dari sampel acak yang digunakan untuk menaksir parameter populasi. Penaksir parameter populasi dapat berupa :

a. Taksiran titik.

Taksiran titik untuk parameter populasi ditentukan dengan menggunakan konsep penaksiran yang telah dijelaskan pada pasal 4.2 dan 4.3.

Contoh.

1. Distribusi Binomial : $X \sim b(x; n, p)$, p tidak diketahui

Di sini : $\theta = p$ yaitu proporsi sukses dalam populasi. Penaksir titik untuk p adalah :

$$\hat{p} = T(X_1, X_2, ..., X_n) = \frac{\sum_{i=1}^{n} X_i}{n},$$

dengan $\sum_{i=1}^{n} X_i$ adalah banyaknya sukses terjadi dalam n ulangan

2. Distribusi Normal : $X \sim N(x; \mu, \sigma^2)$, μ dan σ^2 keduanya tidak diketahui

Di sini $\boldsymbol{\theta} = (\mu, \sigma^2) \in R^2$ yaitu: $\theta_1 = \mu$, rata-rata populasi dan $\theta_2 = \sigma^2$, variansi populasi

Penaksir titik untuk $\mu \rightarrow \hat{\mu} = \hat{\theta}_{mle} = \overline{X}$ yaitu rata-rata sampel acak dan

Penaksir titik untuk $\sigma^2 \rightarrow \hat{\sigma}^2 = s^2$ yaitu variansi sampel acak.

b. Taksiran selang

Namun dalam aplikasinya, sangat diperlukan untuk mendapatkan taksiran yang dapat dinyatakan dalam suatu selang atau interval, di mana parameter populasi yang akan ditaksir diyakini terletak dalam selang tersebut. Selang yang demikian disebut selang kepercayaan untuk parameter.

Pengertian selang kepercayaan (*confidence interval*) untuk parameter.

Misalkan statistik $\hat{\Theta}$ merupakan penaksir takbias bagi parameter θ, maka taksiran selang adalah selang yang berbentuk $\hat{\theta}_1 < \theta < \hat{\theta}_2$; dengan $\hat{\theta}_1$ dan $\hat{\theta}_2$ tergantung pada nilai statistik $\hat{\Theta}$ untuk suatu sampel $\hat{\Theta}$. Artinya, jika sampel yang digunakan berbeda, nilai-nilai $\hat{\theta}_1$ dan $\hat{\theta}_2$ juga berbeda. Dari distribusi sampel $\hat{\Theta}$ dapat ditentukan $\hat{\theta}_1$ dan $\hat{\theta}_2$ sedemikian rupa sehingga $P(\hat{\theta}_1 < \theta < \hat{\theta}_2)$ sama dengan setiap nilai yang kita inginkan. Misalnya $\hat{\theta}_1$ dan $\hat{\theta}_2$ dicari sehingga memenuhi $P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 0{,}95$; maka dengan peluang 0,95 sampel acak yang diambil akan menghasilkan suatu selang yang mengandung θ. Selang ini dihitung berdasarkan sampel acak yang diambil, disebut selang kepercayaan 95%. Dengan perkataan lain, kita percaya 95% bahwa selang yang kita hitung mengandung parameter yang sesungguhnya dari populasi.

Umumnya distribusi $\hat{\Theta}$ memungkinkan menghitung k sehingga

$$P(\hat{\Theta} - k < \theta < \hat{\Theta} + k) = 1-\alpha \qquad 0 < \alpha < 1$$

yang disebut selang kepercayaan $(1-\alpha)100\%$ (*confidence interval*).

Jika $\alpha = 0{,}05$, diperoleh selang kepercayaan 95%, dan bila $\alpha = 0{,}01$, diperoleh selang kepercayaan 99% yang lebih lebar. Pecahan $1 - \alpha$ disebut koefisien kepercayaan, dan titik ujung $\hat{\Theta} - k$ dan $\hat{\Theta} + k$ disebut batas selang kepercayaan.

4.5. Selang Kepercayaan Parameter Rata-rata Populasi

Konsep selang kepercayaan untuk parameter rata-rata populasi terdiri dari 6 (enam) kasus yaitu:

a. Selang kepercayaan untuk rata-rata suatu populasi ukuran sampel besar ($n \geq 30$) dengan standar deviasi populasi σ diketahui.
b. Selang kepercayaan untuk rata-rata suatu populasi ukuran sampel kecil ($n < 30$) dengan standar deviasi populasi σ tidak diketahui.
c. Selang kepercayaan untuk selisih rata-rata dua populasi bebas dengan ukuran sampel dari kedua popolasi besar yaitu $n_1 \geq 30$ dan $n_2 \geq 30$.

d. Selang kepercayaan untuk selisih rata-rata dua populasi bebas dengan ukuran sampel dari kedua popolasi kecil yaitu $n_1 < 30$ dan $n_2 < 30$. Asumsi variansi kedua populasi sama.
e. Selang kepercayaan untuk selisih rata-rata dua populasi bebas dengan ukuran sampel dari kedua popolasi kecil yaitu $n_1 < 30$ dan $n_2 < 30$. Asumsi variansi kedua populasi tidak sama.
f. Selang kepercayaan selisih rata-rata untuk pengamatan berpasangan atau tidak bebas dengan ukuran sampel kecil.

4.5.1. Selang kepercayaan $(1 - \alpha)100\%$ rata-rata suatu populasi ($n \geq 30$)

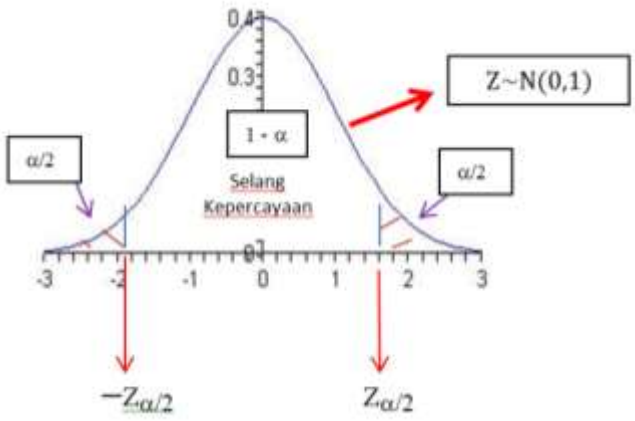
Konsep selang kepercayaan $(1 - \alpha)100\%$ untuk rata-rata populasi (μ) dengan ukuran sampel besar ($n \geq 30$) dan σ diketahui, dimulai dengan informasi distribusi statistik penaksir tak bias untuk μ yaitu rataan sampel acak iid yaitu $\overline{X}$. Telah disebutkan bahwa jika populasi berdistribusi normal atau hampir normal dengan rataan μ dan variansi σ^2, maka distribusi $\overline{X}$ adalah juga normal dengan rata-rata μ dan variansi σ^2/n; sehingga,

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \quad \text{berdistribusi } N(0, 1)$$

Sekarang kita akan menurunkan rumusan selang kepercayaan yang memenuhi bahwa,

$$P\left(\hat{\Theta} - k < \theta < \hat{\Theta} - k\right) = 1 - \alpha,\ 0 < \alpha < 1$$

Dengan $\hat{\Theta}$ di sini adalah $Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$. Kita akan ilustrasikan melalui Gambar 4.5.1.1 berikut ini.



Gambar 4.5.1 Selang kepercayaan $(1 - \alpha)100\%$ rata-rata suatu populasi ($n \geq 30$)

Pada Gambar 4.5.1 di atas diilustrasikan konsep selang kepercayaan sebagai berikut.

$$P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1-\alpha$$

Di mana ,

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

Selang dalam peluang, nyatakan dalam selang:
…. < μ < …..
Sebab yang akan ditentukan adalah selang kepercayaan untuk μ.

Sehingga,

$$P\left(-z_{\alpha/2} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha$$

Atau: $$P\left(\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Maka selang kepercayaan $(1-\alpha)100\%$ untuk rata-rata populasi μ adalah,

$$\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

dengan $\overline{X}$ menyatakan rata-rata sampel ukuran n dari populasi dengan standar deviasi σ yang diketahui. Jika standar deviasi populasi σ tidak diketahui, asalkan $n \geq 30$, σ dapat diganti dengan standar deviasi sampel acak S. Kemudian $z_{\alpha/2}$ menyatakan nilai distribusi normal baku sehingga daerah di sebelah kanannya mempunyai luas $\alpha/2$.

Contoh 4.5.1

Diketahui bahwa rata-rata dan standar deviasi nilai Statistika sampel acak 40 mahasiswa Teknik Mesin ITI, masing-masing 2,8 dan 0,25.

a. Tentukan selang kepercayaan 95% untuk rata-rata populasi, yaitu rata-rata nilai Statistika semua mahasiswa tingkat sarjana.
b. Tentukan selang kepercayaan 99% untuk rata-rata nilai Statistika semua mahasiswa tingkat sarjana.

Penyelesaian.

Taksiran titik untuk μ ialah $\overline{x} = 2{,}8$. Dari soal σ tidak diketahui dan karena ukuran sampelnya besar, standar deviasi σ dapat dihampiri dengan standar deviasi sampel s = 0,25. Maka :

a. Selang kepercayaan 95% $\Rightarrow 1-\alpha = 95\% \Rightarrow \alpha = 5\% = 0{,}05 \Rightarrow \alpha/2 = 0{,}025$

$$2{,}8 - z_{0{,}025}\frac{0{,}25}{\sqrt{40}} < \mu < 2{,}8 + z_{0{,}025}\frac{0{,}25}{\sqrt{40}}$$

Dalam Pasal 3, kita telah membahas nilai z-*score* , $z_{0{,}025} = 1{,}96$. Maka :

$$2{,}8 - 1{,}96\ .\ \frac{0{,}25}{\sqrt{40}} < \mu < 2{,}8 + 1{,}96\ .\ \frac{0{,}25}{\sqrt{40}} \qquad \text{atau : } 2{,}72 < \mu < 2{,}88$$

b. Selang kepercayaan 99% $\Rightarrow 1-\alpha = 99\% \Rightarrow \alpha = 1\% = 0{,}01 \Rightarrow \alpha/2 = 0{,}005$

$$2{,}8 - z_{0{,}005}\frac{0{,}25}{\sqrt{40}} < \mu < 2{,}8 + z_{0{,}005}\frac{0{,}25}{\sqrt{40}}$$

Dengan menggunakan Program Minitab seperti yang telah dijelaskan sebelumnya diperoleh nilai $z_{0{,}005}$ seperti yang diberikan pada Tabel 4.5.5.1.

Tabel 4.5.1 Nilai $z_{0{,}005}$

P(X ≤ x)	x
0,995	2,57583

Jadi: $z_{0{,}005} = 2{,}576$.

Maka selang kepercayaan 99% adalah

$$2{,}8 - 2{,}576\ .\ \frac{0{,}25}{\sqrt{40}} < \mu < 2{,}8 + 2{,}576\ .\ \frac{0{,}25}{\sqrt{40}} \quad \text{atau : } 2{,}698 < \mu < 2{,}902$$

Dari a dan b terlihat bahwa semakin tinggi tingkat kepercayaan maka diperoleh selang kepercayaan yang lebih lebar.

4.5.2. Selang kepercayaan (1 – α)100% rata-rata suatu populasi (n < 30)

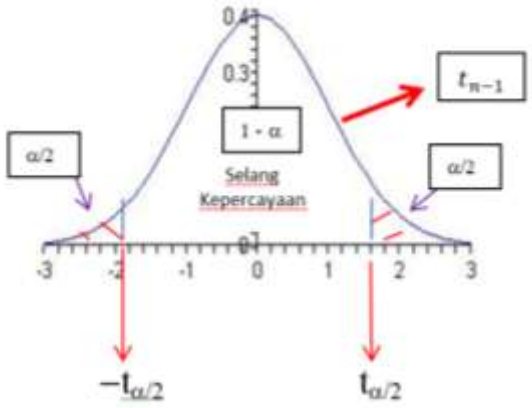
Konsep selang kepercayaan (1 – α)100% untuk rata-rata populasi (μ) dengan ukuran sampel kecil (n < 30) dan σ tidak diketahui tidak lagi menggunakan distribusi Z. Pada halaman 49 telah disebutkan bahwa jika populasi berdistribusi normal atau hampir normal dengan rataan μ dan standar deviasi σ namun tidak diketahui, maka distribusi $\overline{X}$ tidak lagi berdistribusi normal, tetapi berdistribusi T di mana,

$T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ berdistribusi T dengan derajat kebebasan $\nu = n - 1$

Sekarang kita akan menurunkan rumusan selang kepercayaan yang memenuhi bahwa,

$$P\left(\hat{\Theta} - k < \theta < \hat{\Theta} - k\right) = 1-\alpha,\ 0 < \alpha < 1$$

Dengan $\hat{\Theta}$ di sini adalah $Z = \frac{\bar{X}-\mu}{S/\sqrt{n}}$. Kita akan ilustrasikan melalui Gambar 4.5.2 berikut ini.



Gambar 4.5.2 Selang kepercayaan $(1 - \alpha)100\%$ rata-rata suatu populasi ($n < 30$)

Pada Gambar 4.5.2 di atas diilustrasikan konsep selang kepercayaan sebagai berikut.

$$P\left(-t_{\alpha/2} < T < t_{\alpha/2}\right) = 1-\alpha$$

Di mana ,

$$T = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

Sehingga,

$$P\left(-t_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < t_{\alpha/2}\right) = 1-\alpha$$

Atau: $P\left(\bar{X} - t_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$

Maka selang kepercayaan $(1 - \alpha)100\%$ untuk μ jika σ tidak diketahui dan $n < 30$, ialah,

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

dengan $\overline{X}$ dan S masing-masing menyatakan rataan dan standar deviasi sampel ukuran n < 30 yang diambil dari populasi yang hampir normal, dan $t_{\alpha/2}$ menyatakan nilai dari distribusi t dengan derajat kebebasan $\nu = n - 1$, sehingga daerah di sebelah kanannya seluas $\alpha/2$.

Contoh 4.5.2

Di sebuah depot air minum isi ulang dilakukan percobaan untuk menaksir rata-rata volume galon air (dalam liter). Percobaan dilakukan terhadap 7 galon air dan volumenya adalah 19,1 ; 19,2 ; 19,4 ; 19,1 ; 18,9 ; 19,2 dan 19,0 liter. Tentukan selang kepercayaan 95% untuk rata-rata populasi, yaitu rata-rata volume semua galon air pada depot tersebut bila distribusinya dianggap hampir normal.

Penyelesaian

Dengan menggunakan Program Minitab, kita tentukan dahulu rata-rata dan standar deviasi sampel di atas (Tabel 4.5.2.1).

Tabel 4.5.2.1 Statistik data Volume Galon

Variabel	Rata-rata	Standar Deviasi
Volume Galon	19.129	0.160

Maka diperoleh: $\bar{x}$ = 19,129 dan s = 0,16.

Karena yang diinginkan selang kepercayaan 95%, maka α = 0,05.

Dengan menggunakan Program Minitab, untuk $\nu = 7 - 1 = 6$ (Tabel 4.5.2.2).

Tabel 4.5.2.2 Nilai $t_{0,025,\nu=7}$

P(X ≤ x)	x
0.975	2.44691

Jadi: $t_{0,025,\nu=7} = 2,447$.

Maka selang kepercayaan 95% untuk μ adalah

$$19,129 - 2,447 \cdot \frac{0,160}{\sqrt{7}} < \mu < 19,129 + 2,447 \cdot \frac{0,160}{\sqrt{7}} \quad \text{atau} \quad 18,980 < \mu < 19,277$$

4.5.3. Selang kepercayaan (1 – α)100% selisih rata-rata dua populasi bebas, $n_1 \geq 30$ dan $n_2 \geq 30$

Konsep selang kepercayaan (1 – α)100% untuk selisih rata-rata dua populasi bebas $(\mu_1 - \mu_2)$ dengan ukuran sampel dari kedua popolasi besar yaitu $n_1 \geq 30$ dan $n_2 \geq 30$ diturunkan sebagai berikut. Tinjau dua populasi bebas dengan distribusi sampel rataan masing-masing dinyatakan,

$\overline{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$ dan dengan TLP : $Z_1 = \frac{\overline{X} - \mu_1}{\sigma_1/\sqrt{n_1}} \sim N(0,1)$

dan

$\overline{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$ dan dengan TLP : $Z_2 = \frac{\overline{Y} - \mu_2}{\sigma_2/\sqrt{n_2}} \sim N(0,1)$

Distribusi selisih $Z_1 - Z_2$ juga berdistribusi normal baku dengan rata-rata,

$\mu_1 - \mu_2$

dan variansi (karena saling bebas),

$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

Catatan.

Var(X –Y) = var(X) + Var(Y) – 2 Cov(X, Y), jika saling bebas maka suku Cov(X,Y) = 0.

Jadi distribusi selisih $Z_1 - Z_2$ dapat dituliskan sebagai,

$$Z_1 - Z_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

di mana,

$$Z_1 - Z_2 = \frac{\left(\overline{X} - \overline{Y}\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Sehingga konsep selang kepercayaan dapat diturunkan sebagai berikut,

$$P\left(-z_{\alpha/2} < Z_1 - Z_2 < z_{\alpha/2}\right) = 1 - \alpha$$

Yaitu,

$$P\left(-z_{\alpha/2} < \frac{\left(\overline{X} - \overline{Y}\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < z_{\alpha/2} \right) = 1 - \alpha$$

Atau: $P\left(\left(\overline{X} - \overline{Y}\right) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < \left(\overline{X} - \overline{Y}\right) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right) = 1 - \alpha$

Maka selang kepercayaan $(1 - \alpha)100\%$ untuk rata-rata dua populasi bebas dengan ukuran sampel dari kedua popolasi besar adalah,

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

dengan $\overline{X}$ dan $\overline{Y}$ menyatakan rata-rata sampel acak bebas ukuran n_1 dan n_2 dari populasi yang variansinya masing-masing σ_1^2 dan σ_2^2 (baik n_1 dan $n_2 \geq 30$; jika σ_1^2 dan σ_2^2 tidak diketahui dapat diganti dengan variansi sampel S_1^2 dan S_2^2), dan $z_{\alpha/2}$ adalah nilai kurva normal baku sedimikian rupa sehingga daerah di sebelah kanannya seluas $\alpha/2$.

Contoh 4.5.3

Pada masa pandemi dilaksanakan kuliah Statistika secara daring dengan jumlah peserta yang besar terdiri dari 80 mahasiswa wanita dan 60 mahasiswa pria. Nilai rata-rata Ujian Tengah Semester (UTS) mahasiswa wanita adalah 85 dengan standar deviasi 6, sedangkan nilai rata-rata UTS mahasiswa pria adalah 78 dengan standar deviasi 7. Tentukan selang kepercayaan 96% untuk selisih $\mu_1 - \mu_2$, bila μ_1 menyatakan rata-rata nilai semua mahasiswa wanita dan μ_2 menyatakan rata-rata nilai semua mahasiswa pria yang mungkin mengikuti ujian matakuliah ini.

Penyelesaian

Taksiran titik untuk $\mu_1 - \mu_2$: $\bar{x}-\bar{y} = 85 - 78 = 7$

Karena n_1 dan n_2 keduanya besar, σ_1 dapat diganti $s_1 = 6$ dan σ_2 dengan $s_2 = 7$.

Selang kepercayaan 96%, $\Rightarrow 1 - \alpha = 96\% \Rightarrow \alpha = 4\% = 0{,}04 \Rightarrow \alpha/2 = 0{,}02$

Dengan menggunakan Program Minitab (Tabel 4.5.3).

Tabel 4.5.3 Nilai $z_{0,02}$

P(X ≤ x)	x
0,98	2,05375

Jadi: $z_{0,02} = 2{,}054$

Maka selang kepercayaan 96% untuk $\mu_1 - \mu_2$ adalah,

$$7-(2{,}054)\sqrt{\frac{36}{80}+\frac{49}{60}} < \mu_1-\mu_2 < 7+(2{,}054)\sqrt{\frac{36}{80}+\frac{49}{60}}$$

Atau: $4{,}69 < \mu_1 - \mu_2 < 9{,}31$

4.5.4. Selang kepercayaan (1 – α)100% selisih rata-rata dua populasi bebas, $n_1 < 30$ dan $n_2 < 30$. Asumsi $\sigma_1^2 = \sigma_2^2$ tapi tidak diketahui

Konsep selang kepercayaaan (1 – α)100% untuk selisih rata-rata dua populasi bebas $(\mu_1 - \mu_2)$ dimulai dengan menentukan penaksir tak bias untuk variansi kedua populasi. Di sini $\sigma_1^2 = \sigma_2^2$. Sebut $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Dari,

$$\overline{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right) \quad \text{dan} \quad \overline{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

Maka:

$$\text{Var}\left(\overline{X} - \overline{Y}\right) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

Sehingga: $E\left[S_1^2\right] = \sigma^2$ dan $E\left[S_2^2\right] = \sigma^2$ yang bermakna S_1^2 dan S_2^2 masing-masing penaksir tak bias untuk σ^2.

Maka menurut TLP halaman 48,

$$\frac{\left(\overline{X} - \overline{Y}\right) - \left(\mu_1 - \mu_2\right)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1)$$

Penaksir tak bias untuk σ^2 adalah,

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

Karena,

$$E\left[S_p^2\right] = E\left[\frac{(n_1 - 1)S_1^{\,2} + (n_2 - 1)S_2^{\,2}}{n_1 + n_2 - 2}\right] = \frac{(n_1 - 1)E\left[S_1^{\,2}\right] + (n_2 - 1)E\left[S_2^{\,2}\right]}{n_1 + n_2 - 2}$$

$$= \frac{(n_1 - 1).\sigma^2 + (n_2 - 1).\sigma^2}{n_1 + n_2 - 2} = \sigma^2$$

Sekarang akan ditentukan distribusi dari,

$$\frac{\left(\overline{X} - \overline{Y}\right) - \left(\mu_1 - \mu_2\right)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Dari halaman 49 ,

$$\frac{(n_1-1)S_1^2}{\sigma_1^2} \text{ berdistribusi} \chi^2_{n_1-1} \text{ dan } \frac{(n_2-1)S_2^2}{\sigma_2^2} \text{ berdistribusi} \chi^2_{n_2-1}$$

Dalam konsep teori peluang:

1. Jika

$$\frac{(n_1-1)S_1^2}{\sigma_1^2} \text{ dan } \frac{(n_2-1)S_2^2}{\sigma_2^2} \text{ saling bebas}$$

Maka distribusi : $\frac{(n_1-1)S_1^2}{\sigma_1^2} + \frac{(n_2-1)S_2^2}{\sigma_2^2}$ adalah $\chi^2_{n_1+n_2-1}$

2. S_1^2 saling bebas dengan $\overline{X}$ dan S_2^2 saling bebas dengan $\overline{Y}$.

Dari kedua hal ini dan berdasarkan konsep distribusi T maka,

$$T = \frac{\left[(\overline{X}-\overline{Y})-(\mu_1-\mu_2)\right] / \sqrt{\sigma\left(n_1^{-1}+n_2^{-1}\right)}}{\sqrt{(n_1+n_2-2)S_p^2 / (n_1+n_2-2)\,\sigma^2}} = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_p\left(n_1^{-1}+n_2^{-1}\right)}$$

berdistribusi T dengan derajat bebas,

$$\nu = n_1 + n_2 - 2$$

Maka sekarang kita dapat menurunkan konsep selang kepercayaan $(1-\alpha)100\%$ untuk $\mu_1 - \mu_2$,

$$P\left[-t_{\alpha/2} < \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_p\left(n_1^{-1}+n_2^{-1}\right)} < t_{\alpha/2}\right] = 1-\alpha$$

Atau,

$$P\left[(\overline{X}-\overline{Y})-t_{\alpha/2}\,S_p\left(n_1^{-1}+n_2^{-1}\right) < \mu_1-\mu_2 < (\overline{X}-\overline{Y})+t_{\alpha/2}\,S_p\left(n_1^{-1}+n_2^{-1}\right)\right] = 1-\alpha$$

Maka selang kepercayaan $(1-\alpha)100\%$ untuk $\mu_1 - \mu_2$,

$$(\vec{x}-\bar{y})-t_{\alpha/2}\,S_p\left(n_1^{-1}+n_2^{-1}\right) < \mu_1-\mu_2 < (\vec{x}-\bar{y})+t_{\alpha/2}\,S_p\left(n_1^{-1}+n_2^{-1}\right)$$

dengan $\overline{X}$ dan $\overline{Y}$ menyatakan rata-rata sampel kecil, masing-masing dengan ukuran n_1 dan n_2 dari populasi yang berdistribusi hampir normal, S_p menyatakan standar deviasi gabungan, yaitu

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

dengan S_1^2 dan S_2^2 masing-masing menyatakan variansi sampel, dan $t_{\alpha/2}$ adalah nilai pada distribusi t dengan derajat kebebasan $\nu = n_1 + n_2 - 2$, sehingga daerah di sebelah kanannya seluas $\alpha/2$.

Contoh 4.5.4

Dalam sebuah pelaksanaan perkuliahan secara daring dengan jumlah peserta yang kecil, hendak dibandingkan rata-rata nilai UTS matakuliah Matematika I dengan rata-rata nilai UTS matakuliah Statistika. Nilai rata-rata UTS 18 mahasiswa peserta Matematika I adalah 82 dengan standar deviasi 5, sedangkan nilai rata-rata UTS 15 mahasiswa peserta Statistika adalah 75 dengan standar deviasi 4. Tentukan selang kepercayaan 96% untuk selisih rata-rata kedua populasi bila dianggap kedua populasi berdistribusi hampir normal dengan variansi yang sama tetapi tidak diketahui.

Penyelesaian

Misalkan μ_1 dan μ_2 masing-masing menyatakan rata-rata populasi semua peserta yang mungkin untuk kedua matakuliah . Maka taksiran titik untuk $\mu_1 - \mu_2$ adalah

$$\bar{x}_1 - \bar{x}_2 = 82 - 75 = 7$$

Berikutnya, kita tentukan dahulu :

$$s_p^2 = \frac{(18-1)(25)+(15-1)(16)}{18+15-2} = 20{,}94 \quad \Rightarrow \; s_p = 4{,}58$$

Selang kepercayaan 96%, $\Rightarrow 1 - \alpha = 96\% \Rightarrow \alpha = 4\% = 0{,}04 \Rightarrow \alpha/2 = 0{,}02$

Dengan menggunakan Program Minitab, $t_{0,02}$ dengan derajat kebebasan $\nu = 18 + 15 - 2 = 31$ (Tabel 4.5.4).

Tabel 4.5.4 Nilai $t_{0,02,\ \nu=31}$

$P(X \le x)$	x
0,98	2,14383

Jadi: $t_{0,02\ ,\ \nu = 31} = 2{,}14$

Maka selang kepercayaan 96% adalah

$$7 - (2{,}14)(4{,}58)\sqrt{\frac{1}{18}+\frac{1}{15}} < \mu_1 - \mu_2 < 7 + (2{,}14)(4{,}58)\sqrt{\frac{1}{18}+\frac{1}{15}}$$

Atau : $3{,}57 < \mu_1 - \mu_2 < 10{,}43$.

4.5.5. Selang kepercayaan (1 – α)100% selisih rataan dua populasi bebas, $n_1 < 30$ dan $n_2 < 30$. Asumsi $\sigma_1^2 \neq \sigma_2^2$ dan tidak diketahui

Konsep selang kepercayaaan (1 – α)100% dimulai dengan asumsi bahwa $\sigma_1^2 \neq \sigma_2^2$ dan tidak diketahui. Dalam kasus ini digunakan Persamaan Welch-Satterthwaite (Persamaan 4.5.5) untuk menentukan derajat bebas bagi kombinasi linear dari variansi sampel bebas,

$$\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$$

Distribusi,

$$T = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

adalah berdistribusi T dengan derajat bebas,

$$\nu = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]} \qquad (4.5.5)$$

Maka selang kepercayaan (1 – α)100% untuk selisih rata-rata $\mu_1 - \mu_2$, dengan ukuran sampel kecil, dan $\sigma_1^2 \neq \sigma_2^2$ tapi tidak diketahui, ialah,

$$(\overline{x}_1 - \overline{x}_2) - t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{x}_1 - \overline{x}_2) + t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

dengan $\overline{X}$ dan S_1^2 dan $\overline{Y}$ dan S_2^2 menyatakan rata-rata dan variansi sampel kecil bebas berukuran masing-masing populasi dengan ukuran sampel n_1 dan n_2 berdistribusi normal atau hampir normal, dan $t_{\alpha/2}$ nilai distribusi t dengan derajat kebebasan seperti diberikan oleh Persamaan 4.5.5 sehingga luas di sebelah kanannya $\alpha/2$.

Contoh 4.5.5

Hendak dibandingkan rata-rata berat badan bayi laki-laki dengan rata-rata berat badan bayi perempuan ketika lahir. Rata-rata berat badan dari 15 bayi laki-laki adalah 4 kg dengan standar deviasi 1kg, sedangkan rata-rata berat badan dari 10 bayi perempuan adalah 2,5 kg dengan standar deviasi 0,5 kg. Tentukan selang kepercayaan 95% untuk selisih rata-rata sesungguhnya dari berat badan bayi laki-laki dan perempuan ketika lahir. Anggap bahwa pengamatan berasal dari populasi normal dengan variansi yang berbeda dan tidak diketahui.

Penyelesaian

Misalkan μ_1 dan μ_2 masing-masing menyatakan rata-rata populasi berat badan bayi laki-laki dan perempuan yang sesungguhnya. Maka taksiran titik untuk $\mu_1 - \mu_2$ adalah

$$\bar{x}_1 - \bar{x}_2 = 4 - 2{,}5 = 1{,}5$$

dan

$$\nu = \frac{[(1^2/15) + (0{,}5^2/10)]^2}{[(1^2/15)^2/(15-1)] + [(0{,}5^2/10)^2/(10-1)]} \approx 21$$

Selang kepercayaan 95%, $\Rightarrow 1 - \alpha = 95\% \Rightarrow \alpha = 5\% = 0{,}05 \Rightarrow \alpha/2 = 0{,}025$

Dengan menggunakan Program Minitab diperoleh $t_{0{,}025}$ dengan $\nu = 21$: $t_{0{,}025\ ,\ \nu = 21} = 2{,}08$ (Tabel 4.5.5).

Tabel 4.5.5 Nilai $t_{0{,}025\ ,\ \nu = 21}$

P(X ≤ x)	x
0.975	2.07961

Maka selang kepercayaan 95% adalah

$$1{,}5 - (2{,}08)\sqrt{\frac{1^2}{15} + \frac{0{,}5^2}{10}} < \mu_1 - \mu_2 < 1{,}5 + (2{,}08)\sqrt{\frac{1^2}{15} + \frac{0{,}5^2}{10}}$$

Atau : $0{,}87 < \mu_1 - \mu_2 < 2{,}13$.

4.5.6. Selang kepercayaan (1 – α)100% selisih rata-rata pengamatan berpasangan

Selang kepercayaan selisih rata-rata untuk pengamatan berpasangan atau tidak bebas dengan ukuran sampel kecil dimulai dengan konsep pengamatan berpasangan atau tidak bebas (dependen) yang akan diilustrasikan dalam kasus-kasus berikut.

1. Misalkan akan diproduksi sebuah vitamin baru untuk menambah nafsu makan anak balita supaya berat badannya bertambah. Percobaan dilakukan terhadap tikus. Berat badan tikus ditimbang sebelum diberi makan vitamin, lalu dicatat. Kemudian tikus diberi makan vitamin untuk beberapa waktu lamanya. Setelah itu berat badan tikus dicatat. Dalam kasus ini kita mempunyai data berat badan masing-masing tikus sebelum dan sesudah memakan vitamin penambah nafsu makan tersebut.
2. Kasus kedua adalah masalah perlakuan program penurunan berat badan. Peserta program biasanya diukur berat badannya sebelum dan sesudah menjalani program.

Data pada kedua kasus di atas disebut pengamatan berpasangan atau tidak bebas. Kita mempunyai data "sebelum" dan "sesudah" dari obyek atau subyek yang sama. Namun data berpasangan bisa juga bukan berasal dari obyek atau subyek yang sama seperti yang akan diilustrasikan pada contoh ketiga berikut ini.

3. Misalkan kita ingin membandingkan dua metode pembelajaran yaitu metode luring dan metode daring. 20 orang mahasiswa dibagi dalam 2 kelompok secara acak, artinya tiap mahasiswa mempunyai peluang yang sama untuk masuk ke dalam kelompok satu atau dua. Masing-masing kelompok terdiri dari 10 mahasiswa sehingga kita mempunyai 10 pasang mahasiswa, di mana tiap orang dalam pasangan mempunyai tingkat inteligensi yang hampir sama. Kemudian setiap kelompok diberi materi yang sama, dan setelah beberapa waktu perkuliahan tiap kelompok diberi ujian yang sama.

Data pada contoh ketiga ini juga data berpasangan, meski obyek dan subyeknya tidak sama.

Konsep selang kepercayaan $(1 - \alpha)100\%$ untuk $\mu_1 - \mu_2 = \mu_D$ dengan pengamatan berpasangan sama dengan konsep selang kepercayaan untuk μ dengan ukuran sampel kecil (Pasal 4.4.2), tetapi dalam hal ini yang akan ditentukan adalah selang kepercayaan untuk selisih skor pasangan pengamatan yaitu μ_D.

Selang kepercayaan $(1 - \alpha)100\%$ untuk $\mu_1 - \mu_2 = \mu_D$, pengamatan pasangan, ialah

$$\bar{d} - t_{\alpha/2}\frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{\alpha/2}\frac{s_d}{\sqrt{n}}$$

dengan $\bar{d}$ dan s_d menyatakan rata-rata dan simpangan baku selisih n pasangan pengukuran dan $t_{\alpha/2}$ menyatakan nilai distribusi t dengan derajat kebebasan $\nu = n - 1$, sehingga luas di sebelah kanannya $\alpha/2$.

Contoh 4.5.6.

Duapuluh empat mahasiswa Teknik Mesin ITI dibagi menjadi 12 pasangan, tiap orang dalam pasangan mempunyai tingkat inteligensi yang hampir sama. Seorang dari tiap pasangan dipilih secara acak dan dimasukkan ke dalam kelompok yang diberi kuliah Matematika Teknik secara daring. Anggota lainnya masuk kelompok yang mengikuti kuliah secara luring. Pada akhir semester tiap kelompok diberi ujian yang sama dan nilai mereka diberikan pada Tabel 4.5.6 berikut.

Tabel 4.5.6 Data berpasangan

Pasangan	Daring	Luring	d
1	75	82	–7
2	64	59	5
3	82	87	–5
4	60	70	–10
5	85	90	–5
6	75	80	–5
7	82	90	–8
8	61	65	–4
9	79	87	–8
10	85	88	–3
11	79	86	–7
12	88	92	–4

Tentukan selang kepercayaan 98% untuk selisih sesungguhnya rata-rata nilai kedua cara belajar.

Penyelesaian

Karena pengamatan dipasangkan, maka : $\mu_1 - \mu_2 = \mu_D$, dengan taksiran titik diberikan oleh

$$\bar{d} = \frac{-7 + 5 + ... + (-7) + (-4)}{12} = -5{,}08$$

Dan

$$s_d^2 = \frac{n\sum_{i=1}^{12} d_i^2 - \left(\sum_{i=1}^{12} d_i\right)^2}{n(n-1)} = 14{,}27 \qquad \Rightarrow s_d = \sqrt{14{,}27} = 3{,}78$$

Selang kepercayaan 98%, $\Rightarrow 1 - \alpha = 98\% \Rightarrow \alpha = 2\% = 0{,}02 \Rightarrow \alpha/2 = 0{,}01$

Dengan Program Minitab $t_{0{,}01}$ dengan $\nu = 12 - 1 = 11$: $t_{0{,}01\ ,\ \nu = 11} = 2{,}72$

Maka selang kepercayaan 98% adalah

$$-5{,}08 - (2{,}72)\frac{3{,}78}{\sqrt{12}} < \mu_D < -5{,}08 + (2{,}72)\frac{3{,}78}{\sqrt{12}} \quad \text{atau :} \quad -8{,}05 < \mu_D < -2{,}12.$$

4.6. Selang Kepercayaan Proporsi Populasi

4.6.1. Selang kepercayaan $(1 - \alpha)100\%$ proporsi sebuah populasi

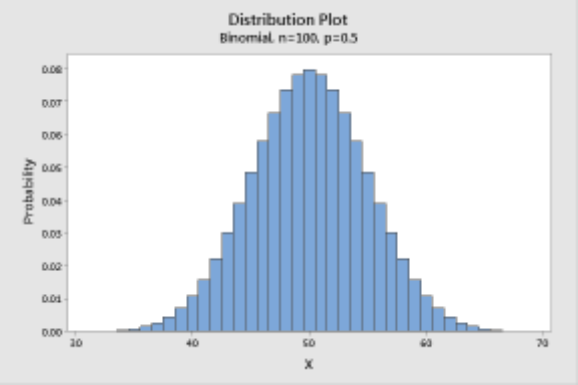
Konsep selang kepercayaan untuk proporsi sebuah populasi (p) dimulai dengan pembahasan tentang distribusi penaksir titik pada distribusi binomial.

Hampiran Normal terhadap Binomial.

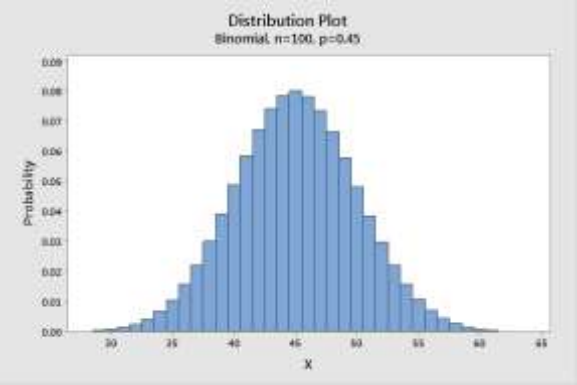
Bila X ~ b(x; n, p), maka menurut konsep TLP halaman 48 maka bentuk limit distribusi,

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \qquad (4.6.1)$$

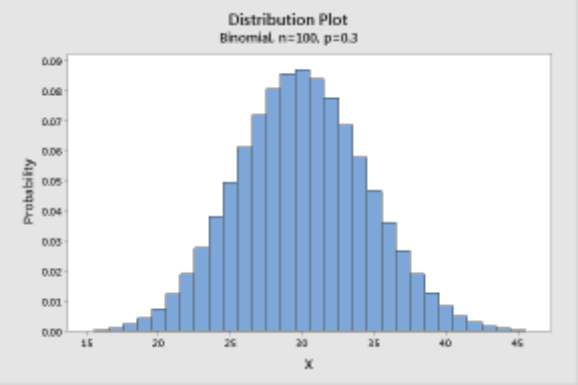
bila n $\to\infty$ adalah distribusi normal baku N(z; 0, 1). Distribusi normal baku memberikan hampiran yang amat baik terhadap distribusi binomial bila n besar dan p dekat ½ seperti yang diilustrasikan pada dua gambar berikut ini, yang merupakan histogram dari Distribusi Binomial dengan n = 100. Di sini penaksir proporsi populasi yaitu $\hat{p}$ yang merupakan proporsi sampel. Pada Gambar 4.6.1.1 nilai p = 0,5 dan pada Gambar 4.6.1.2 nilai p = 0,45. Terlihat bahwa grafik histogram Distribusi Binomial seperti grafik kurva Distribusi Normal dan simetris seperti lonceng. Tetapi jika p tidak dekat dengan ½ maka (Gambar 4.6.1.3 dan 4.6.1.4) grafik histogram Distribusi Binomial hampir seperti grafik kurva Distribusi Normal tetapi tidak simetris seperti lonceng.



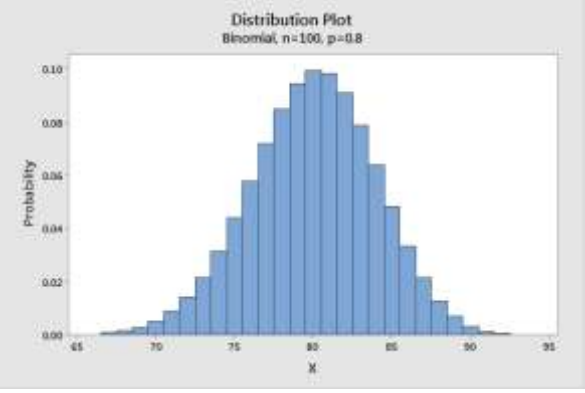
Gambar 4.6.1.1 Grafik Normal n = 100, p=0,5



Gambar 4.6.1.2 Grafik Normal n = 100, p=0,45



Gambar 4.6.1.3 Grafik Normal n = 100, p=0,3



Gambar 4.6.1.4 Grafik Normal n = 100, p=0,8

Berdasarkan Persamaan 4.6.1 maka sekarang kita akan menurunkan selang kepercayaan untuk parameter proporsi pada Distribusi Binomial,

$$P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1-\alpha$$

Di mana ,

$$Z = \frac{X - np}{\sqrt{n\hat{p}(1-\hat{p})}} = \frac{X/n - p}{\sqrt{n\hat{p}(1-\hat{p})}} = \frac{\hat{P} - p}{\sqrt{n\hat{p}(1-\hat{p})}}$$

Sehingga,

$$P\left(-z_{\alpha/2} < \frac{X - np}{\sqrt{n\hat{p}(1-\hat{p})}} < z_{\alpha/2}\right) = 1-\alpha$$

Atau:

$$P\left(-z_{\alpha/2} < \frac{\hat{P} - p}{\sqrt{n\hat{p}(1-\hat{p})}} < z_{\alpha/2}\right) = 1-\alpha$$

Atau: $P\left(\hat{P} - z_{\alpha/2}\sqrt{n\hat{p}(1-\hat{p})} < p < \hat{P} + z_{\alpha/2}\sqrt{n\hat{p}(1-\hat{p})}\right) = 1-\alpha$

Selang kepercayaan untuk parameter binomial p, $n \geq 30$, diberikan oleh,

$$\hat{P} - z_{\alpha/2}\sqrt{n\hat{p}(1-\hat{p})} < p < \hat{P} + z_{\alpha/2}\sqrt{n\hat{p}(1-\hat{p})}$$

dengan $\hat{P} = \frac{X}{n}$ merupakan taksiran titik untuk p yang menyatakan proporsi sukses dalam sampel acak ukuran n, x adalah banyaknya sukses dalam sampel acak ukuran n, dan $z_{\alpha/2}$ menyatakan nilai kurva normal baku sehingga luas di sebelah kanannya $\alpha/2$.

Contoh 4.6.1

Diketahui bahwa dari sampel acak 1000 keluarga di suatu kota besar pada suatu provinsi ditemukan bahwa 650 keluarga menggunakan jaringan nirkabel WiFi di rumahnya. Tentukan selang kepercayaan 95% untuk proporsi sesungguhnya dari keluarga yang menggunakan jaringan nirkabel WiFi di seluruh kota pada provinsi tersebut.

Penyelesaian

Taksiran titik untuk p adalah : $\hat{p} = \frac{650}{1000} = 0{,}65$.

Selang kepercayaan 95%, $\Rightarrow 1 - \alpha = 95\% \Rightarrow \alpha = 5\% = 0{,}05 \Rightarrow \alpha/2 = 0{,}025$ dan $z_{0{,}025} = 1{,}96$

Maka selang kepercayaan 95% adalah

$$0{,}65 - (1{,}96)\sqrt{\frac{(0{,}65)\ (0{,}35)}{1000}} < p < 0{,}65 + (1{,}96)\sqrt{\frac{(0{,}65)\ (0{,}35)}{1000}} \quad = 0{,}62 < p < 0{,}68$$

4.6.2. Selang kepercayaan (1 – α)100% selisih proporsi dua populasi

Tinjau dua populasi bebas yang masing-masing berasal dari peubah Distribusi Binomial yaitu,

$$X \sim b(x; n_1, p_1) \text{ dan } Y \sim b(y; n_2, p_2)$$

maka menurut konsep TLP maka bentuk limit distribusi,

$$Z_1 = \frac{X - n_1 p_1}{\sqrt{n_1 \hat{p}_1 (1 - \hat{p}_1)}} = \frac{\hat{P}_1 - p_1}{\sqrt{\hat{p}_1 (1 - \hat{p}_1)/n_1}} \text{ dan}$$

$$Z_2 = \frac{Y - n_2 p_2}{\sqrt{n_2 \hat{p}_2 (1 - \hat{p}_2)}} = \frac{\hat{P}_2 - p_2}{\sqrt{\hat{p}_2 (1 - \hat{p}_2)/n_2}} \qquad (4.6.2.1)$$

masing-masing berdistribusi Normal Baku N(z; 0, 1), dengan $\hat{P}_1 = X/n_1$ dan $\hat{P}_2 = Y/n_2$ masing-masing penaksir titik bagi p_1 dan p_2.

Penaksir titik bagi $p_1 - p_2$ adalah $\hat{P}_1 - \hat{P}_2$ dengan

Maka $Z_1 - Z_2$ juga mempunyai Distribusi Normal dengan rataan,

$$\mu_{\hat{P}_1 - \hat{P}_2} = p_1 - p_2$$

dan karena kedua populasi saling bebas maka variansi,

$$\sigma^2_{\hat{P}_1 - \hat{P}_2} = \frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}$$

Sehingga,

$$Z = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(p_1 - p_2\right)}{\sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}}} \qquad (4.6.2.2)$$

Berdasarkan Persamaan 4.6.2.2 maka sekarang kita akan menurunkan selang kepercayaan untuk parameter selisih proporsi pada dua peubah acak Distribusi Binomial,

$$P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1 - \alpha$$

Di mana ,

$$Z = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(p_1 - p_2\right)}{\sqrt{\frac{\hat{p}_1 (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 (1 - \hat{p}_2)}{n_2}}}$$

Sehingga,

$$P\left(-z_{\alpha/2} < \frac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(p_1 - p_2\right)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} < z_{\alpha/2} \right) = 1-\alpha$$

Maka selang kepercayaan $(1 - \alpha)100\%$ untuk selisih dua parameter binomial, $p_1 - p_2$ ialah

$$\left(\hat{P}_1 - \hat{P}_2\right) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 - p_2 < \left(\hat{P}_1 - \hat{P}_2\right) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

bila $\hat{P}_1$ dan $\hat{P}_2$ masing-masing menyatakan proporsi sukses dalam sampel acak ukuran n_1 dan n_2 , dengan $\hat{p}_1 = \frac{x}{n_1}$ dan $\hat{p}_2 = \frac{y}{n_2}$, dan $z_{\alpha/2}$ nilai kurva normal baku sehingga luas di sebelah kanannya $\alpha/2$.

Contoh 4.6.2

Diketahui bahwa dalam suatu proses produksi dengan menggunakan mesin A dan mesin B akan terdapat produk yang cacat. Sampel diambil dari produksi mesin A dan mesin B untuk melihat perbedaan proporsi yang cacat dalam produksi. Bila 100 dari 2000 produk yang diproduksi dengan mesin A ternyata cacat, dan 90 dari 3000 produk yang diproduksi dengan mesin B ternyata cacat, tentukan selang kepercayaan 90% untuk selisih sesungguhnya proporsi yang cacat dari kedua mesin.

Penyelesaian

Misalkan p_1 dan p_2 menyatakan proporsi sesungguhnya yang cacat dari mesin A dan mesin B.

Maka

$$\hat{p}_1 = \frac{x}{n_1} = \frac{100}{2000} = 0{,}05 \text{ dan } \hat{p}_2 = \frac{y}{n_2} = \frac{90}{3000} = 0{,}03$$

Dan taksiran titik untuk $p_1 - p_2$ adalah : 0,05 – 0,03 = 0,02.

Selang kepercayaan 90%, $\Rightarrow 1 - \alpha = 90\% \Rightarrow \alpha = 10\% = 0, 10 \Rightarrow \alpha/2 = 0{,}05$. Dengan menggunakan Program Minitab diperoleh $z_{0,05} = 1{,}645$ (Tabel 4.6.2).

Tabel 4.6.2 Nilai $z_{0,05}$

P(X ≤ x)	x
0,95	1,64485

Maka selang kepercayaan 90% adalah

$$0{,}02-(1{,}645)\sqrt{\frac{(0{,}05)(0{,}95)}{2000}+\frac{(0{,}03)(0{,}97)}{3000}}<p_1-p_2<0{,}02+(1{,}645)\sqrt{\frac{(0{,}05)(0{,}95)}{2000}+\frac{(0{,}03)(0{,}97)}{3000}}$$

Atau : $0{,}011 < p_1 - p_2 < 0{,}0295$

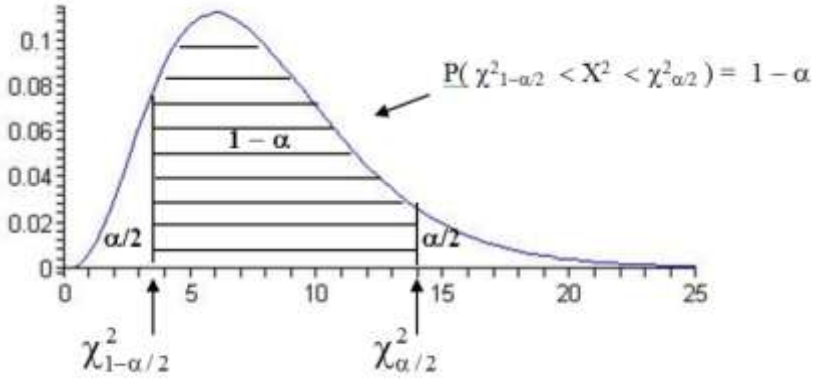
4.7. Selang Kepercayaan Variansi

4.7.1. Selang kepercayaan (1 – α)100% variansi suatu populasi

Taksiran takbias untuk variansi populasi (σ^2) diberikan oleh variansi sampel S^2. Taksiran selang untuk σ^2diturunkan dengan menggunakan statistik

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

yang berdistribusi khi-kuadrat dengan derajat kebebasan ν = n – 1, dan sampel berasal dari populasi normal. Konsep taksiran selang dalam Distribusi Khi-kuadrat ini diilustrasikan dalam Gambar 4.7.1 berikut.



Gambar 4.7.1 Selang kepercayaan (1 – α)100% variansi suatu populasi

$$P\left(\chi^2_{1-\alpha/2}<X^2<\chi^2_{\alpha/2}\right)=1-\alpha$$

Atau,

$$P\left(\chi^2_{1-\alpha/2}<\frac{(n-1)S^2}{\sigma^2}<\chi^2_{\alpha/2}\right)=1-\alpha$$

$\chi^2_{1-\alpha/2}$ dan $\chi^2_{\alpha/2}$ masing-masing menyatakan nilai peubah acak khi-kuadrat dengan derajat kebebasan $\nu = n - 1$, sehingga luas di sebelah kanannya : $1 - \alpha/2$ dan $\alpha/2$.

Maka selang kepercayaan $(1 - \alpha)100\%$ untuk variansi σ^2 suatu populasi normal ialah

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$$

dengan s^2 adalah variansi sampel.

Contoh 4.7.1.

Data berikut menyatakan panjang, dalam cm, 10 bayi perempuan yang baru lahir di sebuah rumahsakit : 49 ; 48 ; 50 ; 49,0 ; 49 ; 51 ; 52 ; 49 ; 49 ; dan 50. Tentukan selang kepercayaan 95% untuk variansi semua bayi yang lahir di rumahsakit tersebut.

Penyelesaian

Pertama-tama kita hitung s^2 : $s^2 = \dfrac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} = 1{,}378$

$$s^2 = \frac{(10)(49^2 + 48^2 + ... + 49^2 + 50^2) - (49 + 48 + ... + 49 + 50)^2}{(10)(10-1)} = 1{,}378$$

Atau, dengan Program Minitab, diperoleh output seperti dberikan pada Tabel 4.7.1 berikut.

Tabel 4.7.1 Statistik panjang badan bayi perempuan

Variabel	Standar deviasi	Variansi
Panjang Badan Bayi Perempuan	1,174	1,378

Selang kepercayaan 95%, $\Rightarrow 1 - \alpha = 95\% \Rightarrow \alpha = 5\% = 0{,}01 \Rightarrow \alpha/2 = 0{,}05$

Berikutnya kita tentukan $\chi^2_{1-\alpha/2}$ dan $\chi^2_{\alpha/2}$ dengan derajat bebas $\nu = 10 - 1 = 9$:

$\chi^2_{1-\alpha/2} = \chi^2_{0{,}975} = 2{,}700$ dan $\chi^2_{\alpha/2} = \chi^2_{0{,}05} = 19{,}023$

Maka selang kepercayaan 95% adalah

$$\frac{(9)(1{,}38)}{19{,}023} < \sigma^2 < \frac{(9)(1{,}38)}{2{,}7} \rightarrow 0{,}65 < \sigma^2 < 4{,}6$$

4.7.2. Selang kepercayaan (1 – α)100% rasio variansi dua populasi

Konsep selang kepercayaan untuk rasio variansi didasarkan pada Distribusi F,

$$F = \frac{U/\nu_1}{V/\nu_2}$$

di mana U dan V dua peubah acak bebas masing-masing berdistribusi khi-kuadrat dengan derajat kebebasan $\nu_1 = n_1 - 1$ dan $\nu_2 = n_2 - 1$ yaitu,

$$X_1^2 = \frac{(n_1 - 1)S_1^2}{\sigma_1^2} \text{ dan } X_2^2 = \frac{(n_2 - 1)S_2^2}{\sigma_2^2}$$

Maka.

$$F = \frac{U/\nu_1}{V/\nu_2} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

Sehingga,

$$P\left(f_{1-\alpha/2} < F < f_{\alpha/2} \right) = 1 - \alpha$$

Atau,

$$P\left(f_{1-\alpha/2} < \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < f_{\alpha/2} \right) = 1 - \alpha$$

$$P\left(f_{1-\alpha/2} < \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < f_{\alpha/2} \right) = 1 - \alpha$$

Selang kepercayaannya akan disederhanakan sebagai berikut.

$$f_{1-\alpha/2} < \frac{S_1^2 \sigma_2^2}{S_2^2 \sigma_1^2} < f_{\alpha/2} \text{ atau } f_{1-\alpha/2} < \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} < f_{\alpha/2}$$

$$\frac{1}{f_{\alpha/2}} < \frac{S_2^2}{S_1^2}\frac{\sigma_1^2}{\sigma_2^2} < \frac{1}{f_{1-\alpha/2}} \text{ atau } \frac{S_1^2}{S_2^2}\frac{1}{f_{\alpha/2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2}\frac{1}{f_{1-\alpha/2}}$$

Di sini: $f_{1-\alpha/2}$ dan $f_{1-\alpha/2}$ mempunyai derajat bebas ν_1 dan ν_2. Kemudian menurut teorema tentang Distribusi F bahwa,

$$f_{1-\alpha/2}(\nu_1, \nu_2) = \frac{1}{f_{\alpha/2}(\nu_2, \nu_1)}$$

Maka selang kepercayaan $(1 - \alpha)100\%$ untuk rasio variansi σ_1^2 / σ_2^2 dapat ditulis,

$$\frac{s_1^2}{s_2^2} \cdot \frac{1}{f_{\alpha/2}(\nu_1, \nu_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} \cdot f_{\alpha/2}(\nu_2, \nu_1)$$

dengan s_1^2 dan s_2^2 menyatakan variansi sampel bebas masing-masing berukuran n_1 dan n_2, diambil dari dua populasi normal, $f_{\alpha/2}(\nu_1, \nu_2)$ menyatakan nilai distribusi F, dengan derajat kebebasan $\nu_1 = n_1 - 1$ dan $\nu_2 = n_2 - 1$, sehingga luas di sebelah kanannya $\alpha/2$, dan $f_{\alpha/2}(\nu_2, \nu_1)$ menyatakan nilai distribusi F, dengan derajat kebebasan $\nu_2 = n_2 - 1$ dan $\nu_1 = n_1 - 1$.

Contoh 4.7.2

Seorang Dosen Penasehat Akademik di suatu perguruan tinggi ingin membandingkan variansi nilai matakuliah Matematika I dan Fisika Dasar untuk mahasiswa semester I. Diketahui bahwa rataan dan variansi dari 20 peserta matakuliah Matematika I masing-masing adalah 80 dan 7. Rataan dan standar deviasi dari 25 peserta matakuliah Fisika Dasar masing-masing adalah 70 dan 5. Tentukan selang kepercayaan 98% untuk $\frac{\sigma_1^2}{\sigma_2^2}$ dan $\frac{\sigma_1}{\sigma_2}$; bila σ_1^2 dan σ_2^2 masing-masing menyatakan variansi populasi nilai matakuliah Matematika I dan Fisika Dasar.

Penyelesaian

Dalam soal ini : $n_1 = 20$, $n_2 = 25$, $s_1 = 7$ dan $s_2 = 5$.

Selang kepercayaan 98%, $\Rightarrow 1 - \alpha = 98\% \Rightarrow \alpha = 2\% = 0{,}02 \Rightarrow \alpha/2 = 0{,}01$

Berikutnya kita tentukan $f_{0,01}(19,24)$ (Tabel 4.7.2) dan $f_{0,01}(24,19)$ (Tabel 4.7.3) dengan derajat bebas $\nu_1 = 20 - 1 = 19$ dan $\nu_2 = 25 - 1 = 24$:

$f_{0,01\,(19,\,24)} = 2{,}76$ dan $f_{0,01\,(24,\,19)} = 2{,}93$

Tabel 4.7.2.1 Nilai $f_{0,01\,(19,\,24)}$

$P(X \leq x)$	x
0,99	2,76239

Tabel 4.7.2.2 Nilai $f_{0,01\,(24,\,19)}$

$P(X \leq x)$	x
0,99	2,92487

Maka selang kepercayaan 98% adalah

$$\frac{49}{25} \cdot \frac{1}{2,76} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{49}{25} \cdot 2,93 \quad \text{atau : } 0,710 < \frac{\sigma_1^2}{\sigma_2^2} < 5,733$$

Selang kepercayaan untuk $\frac{\sigma_1}{\sigma_2}$ dapat diperoleh dengan mengakarkan selang:

$$0,710 < \frac{\sigma_1^2}{\sigma_2^2} < 5,733, \text{ yaitu : } 0,842 < \frac{\sigma_1}{\sigma_2} < 2,394$$

5. Pengujian Hipotesis

5.1. Konsep Pengujian Hipotesis

Definisi. Hipotesis statistik ialah suatu anggapan atau pernyataan, yang mungkin benar atau tidak, mengenai suatu populasi atau lebih.

Anggapan atau pernyataan dalam hipotesis statistik adalah anggapan atau pernyataan tentang parameter populasi.

Hipotesis yang dirumuskan dengan harapan untuk ditolak disebut hipotesis nol dan dinyatakan dengan H_0. Penolakan H_0 menjurus pada penerimaan suatu hipotesis tandingan yang dinyatakan dengan H_1.

Contoh.

Salah satu analisis Statistika yang digunakan oleh penulis dalam penelitian adalah Analisis Regresi. Dalam analisis ini disusun model yang menyatakan hubungan antar beberapa peubah yang diduga memang mempunyai hubungan. Tentu saja harapannya ketika kita mengumpulkan data terkait model tersebut, kita dapat membuktikan bahwa memang terdapat hubungan diantara peubah-peubah tersebut. Maka rumusan hipotesisnya dituliskan sebagai berikut.

H_0: Tidak ada hubungan diantara peubah-peubah

H_1: Ada hubungan diantara peubah-peubah.

Dalam rumusan ini terlihat hipotesis H_0 yang dirumuskan dengan harapan ditolak. Dalam memutuskan apakah kita akan menolak H_0 sesuai dengan harapan di atas, maka kita menghadapi resiko melakukan kesalahan atau galat. Galat dalam prosedur pengujian hipotesis ada dua yang diberikan dalam definisi berikut.

Definisi.

1. Galat jenis I adalah galat yang dilakukan ketika kita memutuskan menolak H_0 padahal sesungguhnya H_0 benar.
2. Galat jenis II adalah galat yang dilakukan ketika kita memutuskan menerima H_0 padahal sesungguhnya H_0 salah.

Keputusan menolak atau menerima H_0 didasarkan atas nilai dari distribusi sampel tertentu seperti pada konsep selang kepercayaan, antara lain Distribusi Normal, Distribusi Khi-kuadrat, Distribusi T dan Distribusi F. Nilai ini disebut nilai kritis yang akan membentuk selang bilangan

yang terdiri dari daerah penerimaan H_0 dan daerah penolakan H_0. Lebih lanjut daerah penolakan disebut daerah kritis.

Peluang melakukan galat jenis I disebut taraf keberartian pengujian dan dinyatakan dengan α yaitu,

$$\alpha = P(H_0 \text{ ditolak} \mid H_0 \text{ benar})$$

Jika α dikaitkan dengan distribusi sampel yang disebutkan diatas, maka yang dimaksud dengan nilai kritis adalah $z_{\alpha/2}$, $t_{\alpha/2}$, $\chi^2_{\alpha/2}$ dan $f_{\alpha/2}$ seperti yang kita temukan dalam pembahasan selang kepercayaan untuk parameter popolasi. Namun dalam konsep selanjutnya pembaca akan menemukan bahwa kita juga akan menggunakan nilai kritis seperti z_α , t_α , χ^2_α dan f_α.

Biasanya nilai peluang α diambil paling besar adalah 0,05. Dalam hal demikian pengujian dikatakan berarti atau signifikan untuk menolak H_0 . Jika nilai peluang $\alpha = 0{,}01$ maka pengujian dikatakan sangat berarti atau sangat signifikan untuk menolak H_0.

Selain peluang α, dalam prosedur pengujian hipotesis ada dua peluang lain yaitu,

$$\beta = P(H_0 \text{ ditolak} \mid H_0 \text{ benar})$$

dan

$$\gamma = 1 - \beta = P(H_0 \text{ ditolak} \mid H_0 \text{ salah})$$

Dalam prosedur pengujian hipotesis tentu saja diharapkan pengujian yang memberikan nilai peluang α dan β yang kecil, sedangkan nilai peluang γ diharapkan tinggi. Semakin tinggi nilai peluang γ semakin baik prosedur pengujian hipotesis yang dilakukan. Selanjutnya peluang γ disebut kuasa uji dari suatu prosedur pengujian hipotesis.

<u>Contoh menghitung α, β dan γ (Kasus Distribusi Binomial)</u>

Misalkan diketahui bahwa di-awal-awal masa pandemi Covid 19 (Maret 2020), sebuah metode penyembuhan dapat membantu merilekskan pernafasan pasien Covid 19. Metode ini dikatakan efektif dengan proporsi pasien yang merasa rileks dengan metode tersebut 30%. Pada bulan Juni 2020 ada pendapat yang mengatakan bahwa ada metode baru yang lebih efektif dari metode pertama. Untuk membuktikan bahwa metode baru memang lebih efektif dari metode lama, dilakukan penelitian berikut ini.

Dengan teknik tertentu ditentukanlah nilai kritis dan daerah kritis sebagai berikut :

Jika lebih dari 8 pasien dari 20 orang pasien yang dirawat merasa lebih rileks dengan metode yang baru, maka disimpulkan bahwa metode baru lebih efektif dari metode sebelumnya.

Di sini nilai kritis adalah 8 dan daerah kritis : $X > 8$,
di mana X = banyak pasien yang merasa rileks.

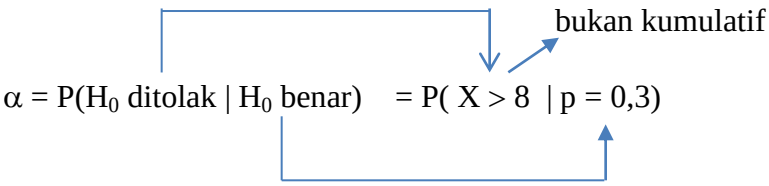
Pemilihan nilai kritis 8 didasarkan pada efektivitas metode lama sebesar 30%, yang artinya jika ada 20 pasien maka dengan metode yang lama paling tidak yang merasa rileks kira-kira 6 orang. Namun menurut pengalaman, dalam prosedur pengujian hipotesis diambil angka tidak tepat 6, tetapi lebih dari 6. Dalam hal ini diambil 8. Nanti nilai kritis ini dapat diatur sedemikian rupa sehingga kita bisa melihat efeknya terhadap nilai-nilai α, β dan γ.

Untuk kasus ini tujuan kita adalah hendak menyimpulkan bahwa metode baru lebih baik dari metode yang lama. Oleh karenanya H_0 (yang disusun untuk ditolak) adalah metode baru sama saja dengan metode lama, artinya proporsi metode baru juga 0,3 lawan alternatif H_1 metode baru lebih baik – yang berarti proporsi metode baru lebih besar dari 0,3. Sehingga rumusan pengujian hipotesisnya adalah sebagai berikut,

$H_0 : p = 0,3$ lawan hipotesis alternatif $H_1 : p > 0,3$, $n = 20$

Di mana p adalah proporsi sesungguhnya pasien yang merasa rileks dengan penyembuhan menggunakan metode baru.

Menentukan α



$= 1 - P(X \leq 8 \mid p = 0,3) = 1 - 0,886669 = 0, 113331$ (termasuk besar)

Catatan.

Dengan menggunakan Program Minitab: Binomial with n = 20 and p = 0.3

x	$P(X \leq x)$
8	0.886669

Misal diambil daerah kritis : $X > 9$. maka

α = P(H_0 ditolak | H_0 benar) = P(X > 9 | p = 0,3)

= 1 – P(X ≤ 9 | p = 0,3) = 1 – 0,952038 =0, 047…. (kecil)

Catatan.

Dengan menggunakan Program Minitab: Binomial with n = 20 and p = 0.3

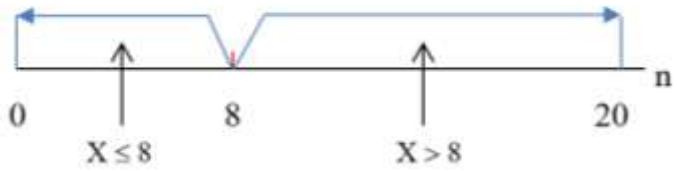
x	P(X ≤ x)
9	0.952038

Kesimpulan. Jika nilai kritis X digeser ke kanan, α semakin kecil.

Jika nilai kritis X digeser ke kiri (paling kecil = 6), α semakin besar.

Menentukan β.

Untuk daerah kritis X > 8, maka daerah penerimaan H_0 adalah X ≤ 8 seperti yang diilustrasikan pada Gambar 5.1.1 berikut.



Daerah penerimaan H_0 Daerah penolakan H_0

Gambar 5.1.1 Daerah penerimaan H_0 dan daerah penolakan H_0 (Binomial)

Maka,

β = P(H_0 diterima | H_0 salah) = P(X ≤ 8 | p > 0,3)

p tidak lagi = 0,3 ; karena p = 0,3 itu di bawah H_0.
Yang dipakai p di bawah H_1 : p > 0,3.
Pilihlah p yang lebih besar 0,3. Misal p = 0,6

Maka,

β = P(X ≤ 8 | p = 0,6) = 0,0565

Catatan.

Dengan menggunakan Program Minitab: Binomial with n = 20 and p = 0.6

x	P(X ≤ x)
8	0,0565264

Jika daerah penerimaan $X \leq 7$: $\beta = P(X \leq 7 \mid p = 0,6) = 0,0210289 \rightarrow$ nilai β semakin kecil.

Catatan.

Dengan menggunakan Program Minitab: Binomial with n = 20 and p = 0.6

x	$P(X \leq x)$
7	0,0210289

Menentukan γ

Untuk daerah penerimaan H_0 $X \leq 8$, maka: $\gamma = 1 - 0,0565 = 0,9435$

Untuk daerah penerimaan H_0 $X \leq 7$, maka: $\gamma = 1 - 0,0210 = 0,979$

Terlihat bahwa jika kita menggeser nilai kritis ke kiri maka nilai peluang γ semakin tinggi.

Contoh menghitung α, β dan γ.

(Kasus Hampiran Distribusi Normal terhadap Distribusi Binomial, ukuran sampel besar)

Misalkan sekarang jumlah pasien ada 100 orang. Jika lebih dari 40 pasien yang dirawat merasa lebih rileks dengan metode yang baru, maka disimpulkan bahwa metode baru lebih efektif dari metode sebelumnya.

Untuk hitung peluang α, β dan γ kita akan menggunakan pendekatan Distribusi Normal terhadap Distribusi Binomial.

Catatan.

Distribusi Binomial : peubah acak diskrit, nilainya bilangan bulat.

Distribusi Normal : peubah acak kontinu, nilainya dalam interval.

Oleh karenanya dalam perhitungan peluang akan ada modifikasi seperti penjelasan berikut.

$H_0 : p = 0,3$ lawan hipotesis alternatif $H_1 : p > 0,3$, n = 100

Menentukan α

bukan kumulatif

$\alpha = P(H_0 \text{ ditolak} \mid H_0 \text{ benar}) = P(X > 40 \mid p = 0,3)$

$= 1 - P(X \leq 40 \mid p = 0,3) \rightarrow$ Selesaikan dengan Distribusi Normal

X = 40 dalam Distribusi Binomial ekivalen dengan selang $39,5 < X < 40,5$ dalam Distribusi Normal

Maka P(X ≤ 40) dalam Distribusi Binomial sama dengan P(X ≤ 40,5) dalam Distribusi Normal. Sehingga :

$\alpha = 1 - P(X < 40,5)$

Untuk menghitung P(X < 40,5) ditentukan dahulu rataan dan variansi Distribusi Normal yaitu,

$\mu = np = 100$, $\sigma^2 = np(1 - p) = 100. (0,3) . (0,7) = 21$, $\sigma = 4,58$

Maka: $\alpha = 1 - P(X < 40,5) = 1 - 0,989064 = 0,01$

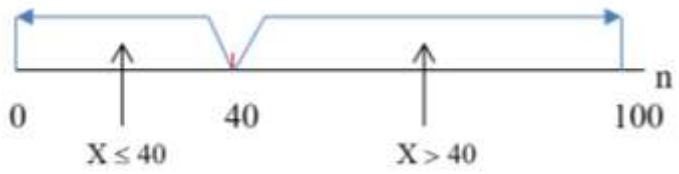
Catatan. Dengan Program Minitab.

P(X ≤ 40,5) = 0,989064

x	P(X ≤ x)
40,5	0,989064

Menentukan β

Untuk daerah kritis X > 40, maka daerah penerimaan H_0 adalah X ≤ 40 seperti yang diilustrasikan pada Gambar 5.1.2 berikut.



Daerah penerimaan H_0 Daerah penolakan H_0

Gambar 5.1.2 Daerah penerimaan H_0 dan daerah penolakan H_0 (Normal)

Seperti telah dijelaskan di atas, X = 40 dalam Distribusi Binomial ekivalen dengan selang 39,5 < X < 40,5 dalam Distribusi Normal, maka:

β = P(H_0 diterima | H_0 salah) = P(X ≤ 40 | p > 0,3)

p tidak lagi = 0,3 ; karena p = 0,3 itu di bawah H_0.
Yang dipakai p di bawah H_1 : p > 0,3.
Pilihlah p yang lebih besar 0,3. Misal p = 0,6

Maka,

β = P(X ≤ 40 | p = 0,6) → dalam Distribusi Binomial

Untuk menghitung P(X < 40,5) ditentukan dahulu rataan dan variansi Distribusi Normal yaitu,

$\mu = np = 60$, $\sigma^2 = npq = 60.\ (0,4) = 24$, $\sigma = 4,9$

Maka dalam Distribusi Normal,

$\beta = P(\ X \leq 40,5) = 0,0000345$

Catatan. Dengan Program Minitab.

$P(X \leq 40,5) = 0,0000345$

x	P(X ≤ x)
40,5	0,0000345

Dan : $\gamma = 1 - \beta = 1 - 0,0000345 = 0,9999$

Dari contoh kasus Distribusi Normal ini kita memperoleh nilai-nilai peluang $\alpha = 0,01$, $\beta = 0,0000345$ dan $\gamma = 0,9999$; yang menunjukkan prosedur pengujian yang baik. Hal ini menunjukkan bahwa dalam Statistika, salah satu cara untuk memperkecil galat adalah dengan memperbesar ukuran sampel n. Semakin besar ukuran sampel, maka statistik sampel semakin representatif untuk menggambarkan parameter populasi yang sebenarnya.

Uji ekaarah dan dwiarah

Sekarang kita akan membahas konsep uji ekaarah dan uji dwiarah yang menjadi dasar pemahaman bagaimana cara menentukan nilai kritis dan daerah kritis pada suatu prosedur pengujian hipotesis. Konsep ini didasarkan pada bentuk hipotesis alternatif H_1.

Misalkan kita akan menguji bahwa nilai parameter $\theta = \theta_0$, yaitu $H_0 : \theta = \theta_0$, dengan hipotesis alternatif adalah salah satu dari $H_1 : \theta < \theta_0$ dan $H_1 : \theta > \theta_0$, yang dituliskan sebagai berikut,

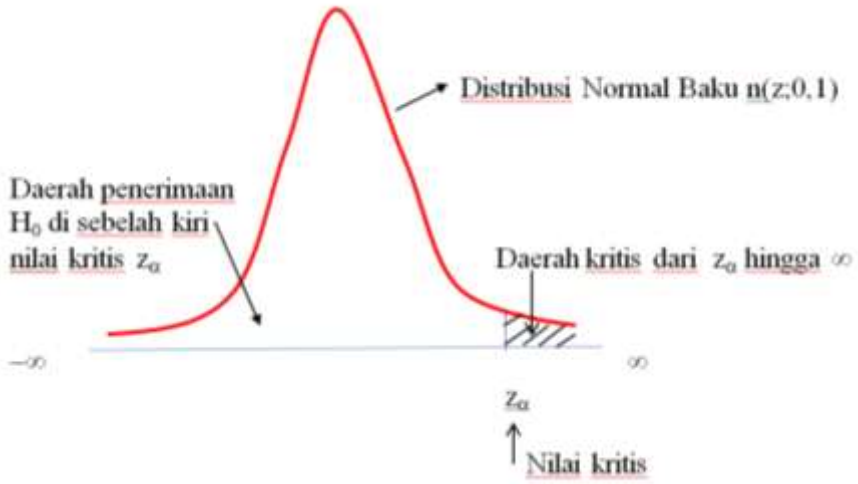
$$H_o : \theta = \theta_o \qquad \text{atau} \qquad H_o : \theta = \theta_o$$
$$H_1 : \theta > \theta_o \qquad\qquad\quad H_1 : \theta < \theta_o$$

Maka prosedur pengujian yang demikian disebut uji ekaarah.

Konsep daerah kritis akan dijelaskan sebagai berikut.
Untuk uji ekaarah berikut,

$$H_o : \theta = \theta_o$$
$$H_1 : \theta > \theta_o$$

seluruh daerah kritis untuk hipotesis tandingan $\theta > \theta_o$ terletak di ujung kanan distribusi sampel seperti yang diilustrasikan pada Gambar 5.1.3 berikut.

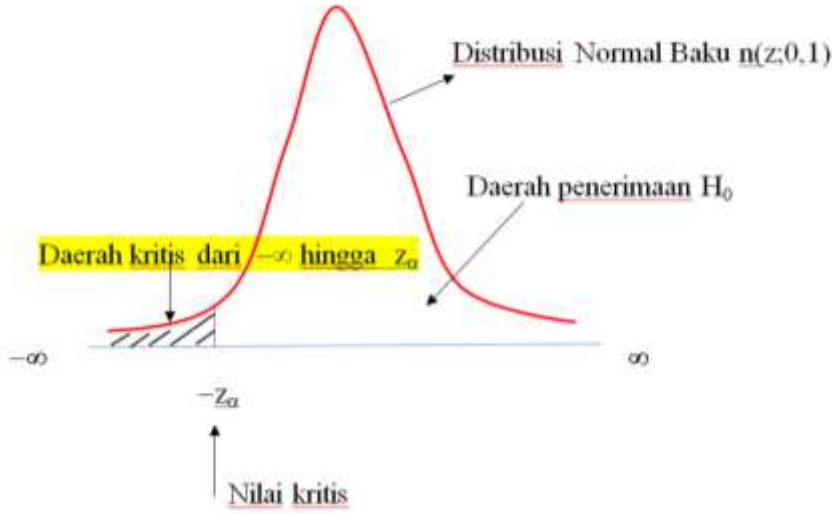


Gambar 5.1.3 Daerah kritis uji ekaarah $\theta > \theta_0$ Distribusi Normal

Untuk uji ekaarah berikut,

$$H_o : \theta = \theta_o$$
$$H_1 : \theta < \theta_o$$

seluruh daerah kritis untuk hipotesis tandingan $\theta > \theta_o$ terletak di ujung kiri distribusi sampel seperti yang diilustrasikan pada Gambar 5.1.4 berikut.



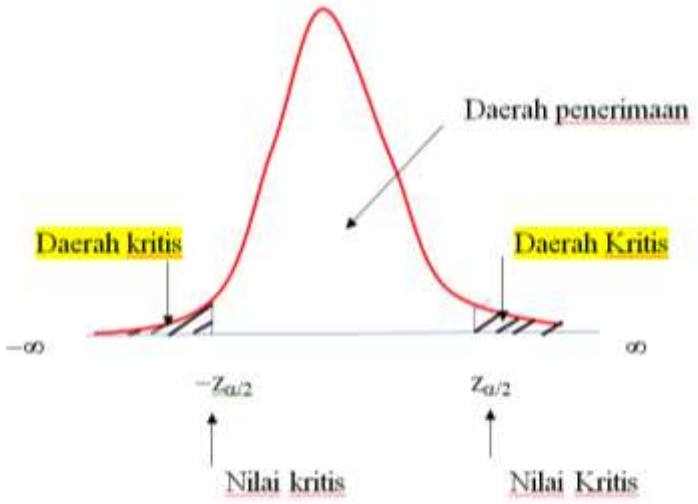
Gambar 5.1.4 Daerah kritis uji ekaarah $\theta < \theta_0$ Distribusi Normal

Sekarang misalkan kita akan menguji $H_0 : \theta = \theta_0$ dengan hipotesis alternatif $H_1 : \theta \neq \theta_0$, yang dituliskan sebagai berikut,

$H_o : \theta = \theta_o$

$H_1 : \theta \neq \theta_o$

maka prosedur pengujian disebut uji dwiarah. Dalam uji dwiarah ini hipotesis alternatif H_1 menyatakan salah satu dari $\theta > \theta_o$ atau $\theta < \theta_o$. seluruh daerah kritis terletak di ujung kiri dan ujung kanan distribusi sampel seperti yang diilustrasikan pada Gambar 5.1.5 berikut.



Gambar 5.1.5 Daerah kritis uji ekaarah $\theta \neq \theta_0$ Distribusi Normal

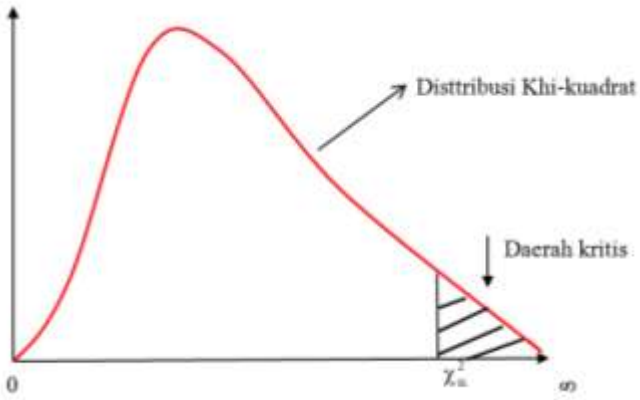
Jika distribusi sampel yang digunakan adalah Distribusi Khi-kuadrat atau Distribusi F, maka konsep daerah kritis dijelaskan sebagai berikut. Penjelasan hanya untuk Distribusi Khi-kuadrat dan berlaku sama dengan Distribusi F.

Untuk uji ekaarah berikut,

$H_o : \theta = \theta_o$

$H_1 : \theta > \theta_o$

daerah kritis di ujung sebelah kanan seperti yang diilustrasikan pada Gambar 5.1.6 berikut.

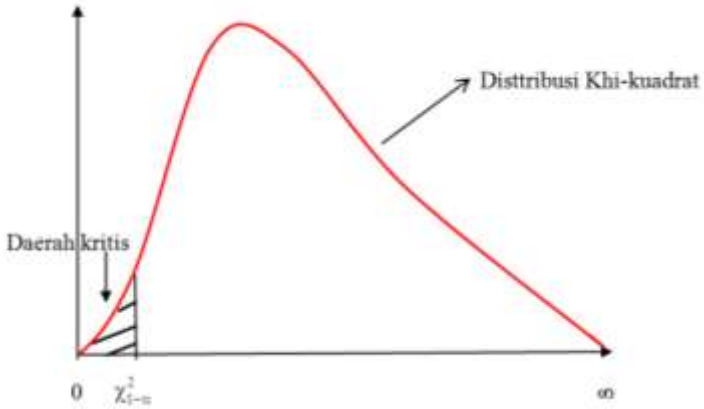


Gambar 5.1.6 Daerah kritis uji ekaarah $\theta > \theta_0$ Distribusi Khi-kuadrat

Untuk uji ekaarah berikut,

$$H_o : \theta = \theta_o$$
$$H_1 : \theta < \theta_o$$

daerah kritis di ujung sebelah kanan seperti yang diilustrasikan pada Gambar berikut.

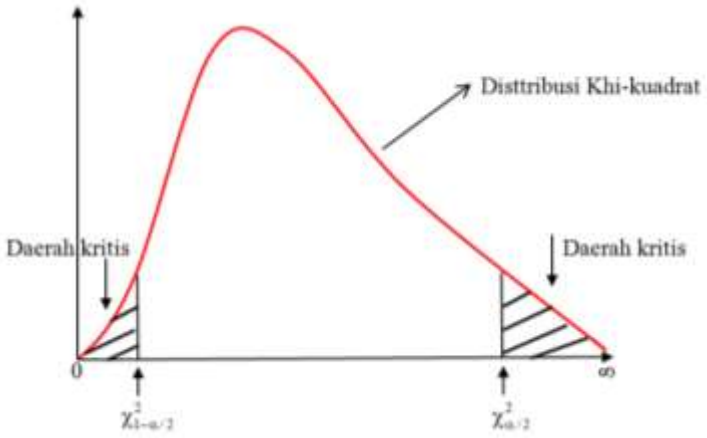


Gambar 5.1.7 Daerah kritis uji ekaarah $\theta < \theta_0$ Distribusi Khi-kuadrat

Untuk uji dwiarah berikut,

$$H_o : \theta = \theta_o$$
$$H_1 : \theta \neq \theta_o$$

daerah kritis di ujung-ujung kiri dan kanan seperti yang diilustrasikan pada Gambar 5.1.8 berikut.



Gambar 5.1.8 Daerah kritis uji ekaarah $\theta \neq \theta_0$ Distribusi Khi-kuadrat

Seperti telah disampaikan sebelumnya bahwa suatu uji dikatakan berarti bila hipotesis nol ditolak pada taraf keberartian 0,05; dan disebut sangat berarti bila hipotesis nol ditolak pada taraf keberartian 0,01.

Selanjutnya kita akan membahas langkah-langkah pengujian hipotesis.

Langkah pengujian hipotesis.

1. Misalkan akan diuji bahwa nilai parameter adalah θ_0. Maka $H_0 : \theta = \theta_0$

   Nilai parameter yang akan diuji di bawah H_0 diberi *subscript* nol.
2. Pilih hipotesis alternatif salah satu dari,

   $H_1 : \theta < \theta_0$ atau $\theta > \theta_0$ atau $\theta \neq \theta_0$
3. Pilih taraf keberartian α
4. Pilih uji statistik yang sesuai dan cari daerah kritis
5. Hitung nilai statistik dari sampel acak ukuran n
6. Kesimpulan :

   Tolak H_0 bila statistik mempunyai nilai dalam daerah kritis. Jika tidak, terima H_0.

5.2. Pengujian Hipotesis Rata-rata dan Variansi

Pada Tabel 5.2.1 diberikan ringkasan pengujian hipotesis yang menyangkut rataan dan variansi.

Tabel 5.2.1 Pengujian hipotesis rata-rata dan variansi

No. (1)	H_0 (2)	Uji Statistik (3)	H_1 (4)	Daerah Kritis (5)
1	$\mu = \mu_0$	$Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$; σ diketahui	$\mu < \mu_0$	$Z < -z_\alpha$
			$\mu > \mu_0$	$Z > z_\alpha$
			$\mu \neq \mu_0$	$Z < -z_{\alpha/2}$ dan $Z > z_{\alpha/2}$
2	$\mu = \mu_0$	$T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$; σ tidak diketahui Ukuran sampel kecil dan $\nu = n - 1$	$\mu < \mu_0$	$T < -t_\alpha$
			$\mu > \mu_0$	$T > t_\alpha$
			$\mu \neq \mu_0$	$T < -t_{\alpha/2}$ dan $T > t_{\alpha/2}$
3	$\mu_1 - \mu_2 = d_0$	$Z = \dfrac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}$ σ_1 dan σ_2 diketahui	$\mu_1 - \mu_2 < d_0$	$Z < -z_\alpha$
			$\mu_1 - \mu_2 > d_0$	$Z > z_\alpha$
			$\mu_1 - \mu_2 \neq d_0$	$Z < -z_{\alpha/2}$ dan $Z > z_{\alpha/2}$
4	$\mu_1 - \mu_2 = d_0$	$T = \dfrac{(\overline{X}_1 - \overline{X}_2) - d_0}{S_p\sqrt{(1/n_1) + (1/n_2)}}$ $\sigma_1 = \sigma_2$ tapi tidak diketahui $\nu = n_1 + n_2 - 2$ $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < d_0$	$T < -t_\alpha$
			$\mu_1 - \mu_2 > d_0$	$T > t_\alpha$
			$\mu_1 - \mu_2 \neq d_0$	$T < -t_{\alpha/2}$ dan $T > t_{\alpha/2}$
5	$\mu_1 - \mu_2 = d_0$	$T = \dfrac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{(S_1^2/n_1) + (S_2^2/n_2)}}$ $\sigma_1 \neq \sigma_2$ dan tidak diketahui $\nu = \dfrac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$	$\mu_1 - \mu_2 < d_0$	$T < -t_\alpha$
			$\mu_1 - \mu_2 > d_0$	$T > t_\alpha$
			$\mu_1 - \mu_2 \neq d_0$	$T < -t_{\alpha/2}$ dan $T > t_{\alpha/2}$
6	$\mu_D = d_0$	$T = \dfrac{\overline{D} - d_0}{S_D/\sqrt{n}}$, $\nu = n - 1$ pengamatan berpasangan	$\mu_D < d_0$	$T < -t_\alpha$
			$\mu_D > d_0$	$T > t_\alpha$
			$\mu_D \neq d_0$	$T < -t_{\alpha/2}$ dan $T > t_{\alpha/2}$
7	$\sigma^2 = \sigma_0^2$	$X^2 = \dfrac{(n - 1)S^2}{\sigma_0^2}$, $\nu = n - 1$	$\sigma^2 < \sigma_0^2$	$X^2 < \chi^2_{1-\alpha}$
			$\sigma^2 > \sigma_0^2$	$X^2 > \chi^2_\alpha$
			$\sigma^2 \neq \sigma_0^2$	$X^2 < \chi^2_{1-\alpha/2}$ dan $X^2 > \chi^2_{\alpha/2}$
8	$\sigma_1^2 = \sigma_2^2$	$F = S_1^2/S_2^2$ $\nu_1 = n_1 - 1$ dan $\nu_2 = n_2 - 1$	$\sigma_1^2 < \sigma_2^2$	$F < f_{1-\alpha}(\nu_1, \nu_2)$
			$\sigma_1^2 > \sigma_2^2$	$F > f_\alpha(\nu_1, \nu_2)$
			$\sigma_1^2 \neq \sigma_2^2$	$F < f_{1-\alpha/2}(\nu_1, \nu_2)$ dan $F > f_{\alpha/2}(\nu_1, \nu_2)$

Kita akan membahas contoh-contoh yang digunakan pada pembahasan selang kepercayaan, namun sekarang dengan prosedur pengujian hipotesisnya. Di sini kita juga akan menggunakan Program Minitab baik untuk selang kepercayaan sekaligus untuk pengujian hipotesis.

5.2.1. Uji hipotesis rata-rata suatu populasi ($n \geq 30$)

Contoh 1. (Contoh 4.5.1 halaman 69)

Diketahui bahwa rata-rata dan standar deviasi nilai Statistika sampel acak 40 mahasiswa Teknik Mesin ITI, masing-masing 2,8 dan 0,25. Menurut pengamatan dosen pengajar matakuliah Statistika, rataan nilai Statistika semua mahasiswa program sarjana adalah 3,5. Ujilah pendapat dosen tersebut dengan hipotesis alternatif bahwa rataan sebenarnya lebih kecil dari 3,5 (karena rataan sampel lebih kecil dari rataan yang dihipotesiskan oleh dosen pengajar tersebut). Gunakan $\alpha = 0{,}05$. Anggap populasi berdistribusi normal.

Penyelesaian.

Taksiran titik untuk μ ialah $\bar{x} = 2{,}8$. Dari soal σ tidak diketahui dan karena ukuran sampelnya besar, standar deviasi σ dapat dihampiri dengan standar deviasi sampel s = 0,25.

Dalam masalah ini, rataan populasi yang akan diuji adalah 3,5. Jadi $\mu_0 = 3{,}5$.

1. $H_0 : \mu = 3{,}5$
2. Karena rataan sampel lebih kecil dari rataan populasi, maka kita memilih :

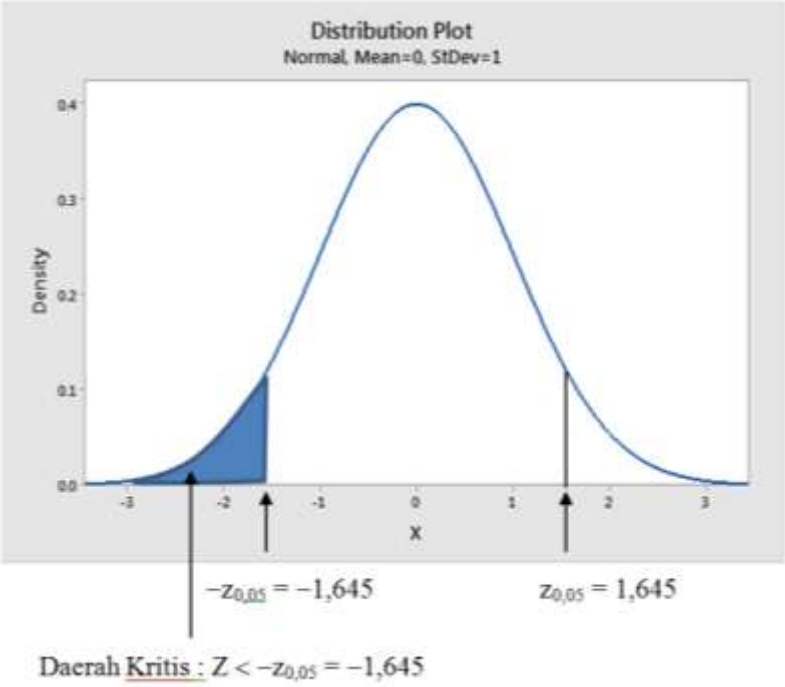
   $H_1 : \mu < \mu_0$, maka $H_1 : \mu < 3{,}5$

   Pilih relasi H_1 berdasarkan kondisi sampel terhadap yang dihipotesiskan
3. Dari soal dipilih $\alpha = 0{,}05$
4. Uji statistik yang sesuai adalah uji nomor 1 pada Tabel 5.2.1,

   $$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$; karena n = 40 > 30, σ dapat diganti s = 0,25

   Dengan $\overline{X}$ adalah rataan sampel, yaitu : $\overline{X} = 2{,}8$.

   Dari Tabel nomor 1 kolom 4 dan 5, berdasarkan H_1 maka daerah kritis adalah ujung sebelah kiri yaitu $Z < -z_\alpha$; yang diilustrasikan pada Gambar 5.2.1.1 berikut.



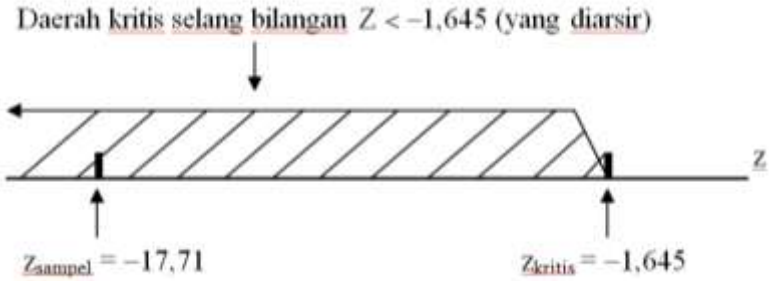
Gambar 5.2.1.1 Daerah kritis Z < –1,645

Maka daerah kritis (daerah penolakan H_0) : Z < –1,645.

5. Perhitungan nilai statistik dari sampel acak ukuran n = 40 :

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{2,8 - 3,5}{0,25 / \sqrt{40}} = -17,71 \rightarrow \text{disebut } z_{hitung} \text{ atau } z_{sampel}$$

6. Kesimpulan dapat dilakukan berdasarkan Gambar 5.2.1.2 berikut. Pada Gambar 5.2.1.2 terlihat bahwa nilai statistik Z (z_{sampel}) terletak dalam daerah kritis. Maka kesimpulannya H_0 ditolak dan H_1 diterima, bahwa berdasarkan sampel ukuran 40 mahasiswa tersebut, rataan nilai Statistika lebih kecil dari 3,5.



Gambar 5.2.1.2 z_{sampel} = 17,71 terletak pada daerah kritis

Cara lain mengambil kesimpulan apakah H_0 ditolak atau diterima adalah dengan menggunakan konsep nilai peluang yang ditulis dengan notasi nilai-p (*p-value*). Nilai-p adalah nilai peluang yang diperoleh dari sampel yang perhitungannya tergantung pada selang bilangan dari daerah kritis. Untuk contoh ini, maka nilai-p adalah,

nilai-p = $P(Z < z_{sampel}) = P(Z < -17{,}71) = 0{,}000$

(Peluang ini ditentukan dengan *Cumulative Probability* pada Program Minitab)

Jika nilai-p < α, maka H_0 ditolak dan jika nilai-p > α, makaH_0 diterima. Untuk contoh ini, karena nilai-p = 0,000 < α = 0,05, maka H_0 ditolak. Pemahamannya α adalah batas maksimal peluang melakukan galat jenis I, sementara berdasarkan sampel diperoleh nilai peluang yang lebih kecil yaitu nilai-p = 0,000 sehingga kita dapat menolak H_0.

Berikut ini diberikan langkah melakukan prosedur pengujian hipotesis untuk Contoh 1 dengan menggunakan Program Minitab.

Langkah pertama pengolahan data dengan menggunakan Program Minitab dimulai dengan menginput data yang akan diolah. Bentuk data dalam contoh-contoh sebelumnya dapat berupa:

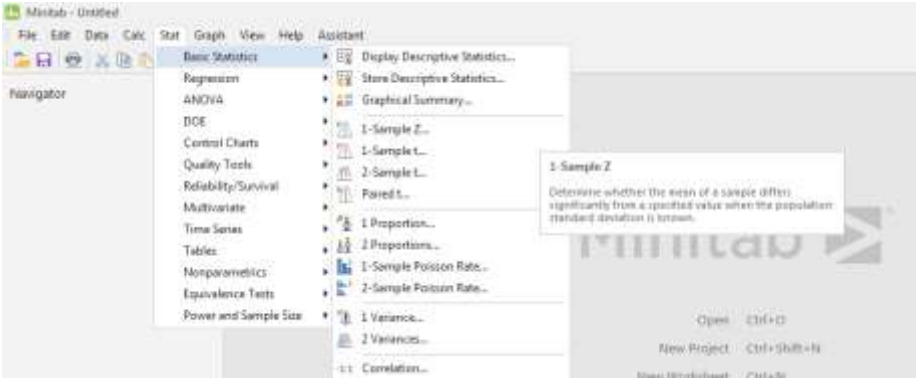
a. Ringkasan data (*Summarized data*)

   Jika data merupakan ringkasan, maka untuk pengolahan data kita bisa langsung klik ikon yang hendak digunakan, seperti ikon *Stat* dalam contoh ini.

b. Data mentah yang harus diinput pada kolom lembar kerja (*worksheet*) Program Minitab (Contoh 4 halaman).

Langkah 1.

Dari ikon *Stat*, pilih *Basic Statistics*, kemudian klik 1-*Sample Z* (Gambar 5.2.1.3).

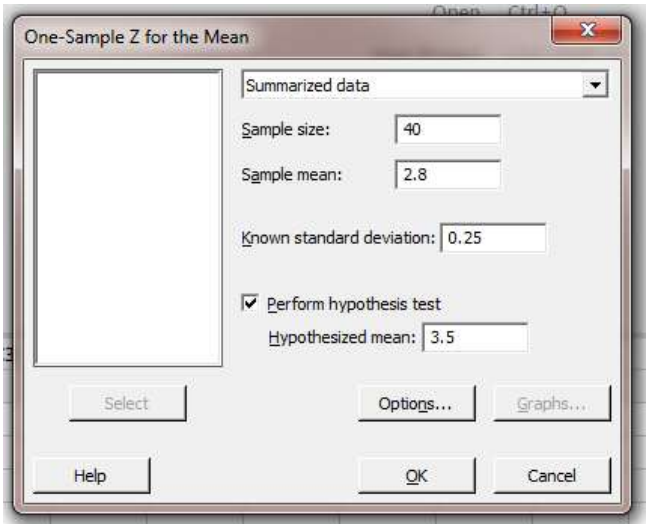


Gambar 5.2.1.3 Menu 1-*Sample Z* Program Minitab

Akan muncul kotak dialog *One-Sample Z for the Mean* seperti pada Gambar 5.2.1.4, kemudian pilih *Summarized Data*. Pilihan ini karena data yang diberikan bukan data mentah, tetapi ringkasan data yaitu ukuran sampel, rata-rata dan standar deviasi.

Gambar 5.2.1.4 Kotak dialog *One-Sample Z for the Mean*

Langkah 2. Setelah diklik *Summarized Data* isikan data-data seperti *Sample size* (ukuran sampel), *Sample mean* (rata-rata sampel) dan *Known standard deviation* (standar deviasi). Kemudian centang *Perform hyphotesis test* dan isikan data *Hypothesized mean* (nilai rataan di bawah H_0). Lalu klik *Options* (Gambar 5.2.1.5).



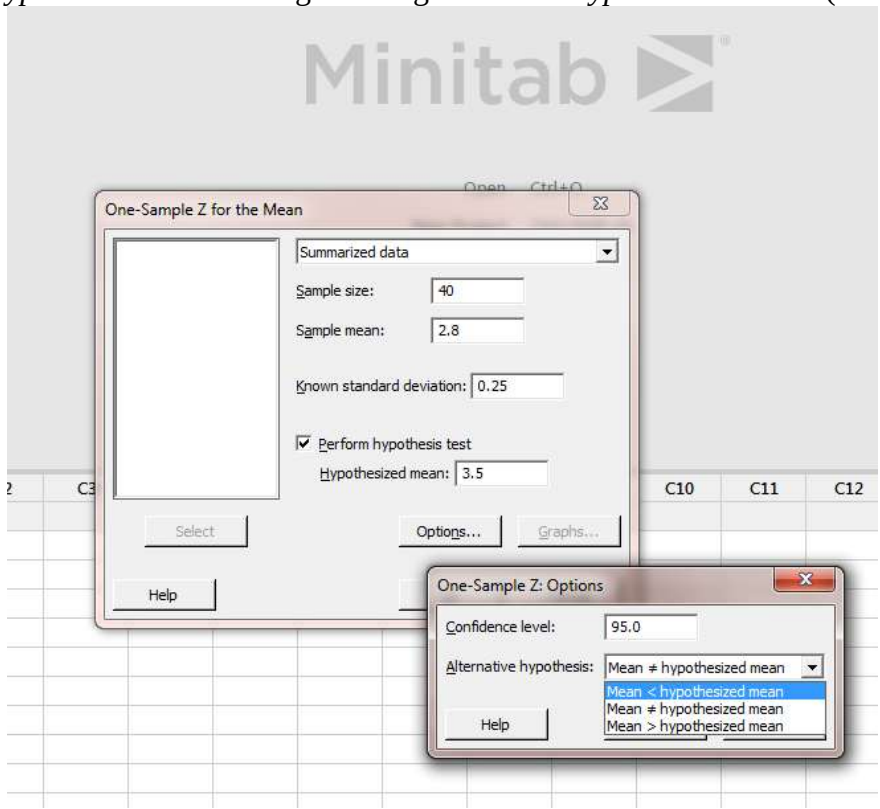
Gambar 5.2.1.5 Langkah 2 Uji *One-Sample Z*

Akan muncul kotak dialog *One-Sample Z* : *Options*. *Default Confidence level* adalah 95%, ini adalah *default* untuk selang kepercayaan 95%. Jadi di sini dapat kita lihat bahwa dalam Program Minitab prosedur penentuan selang kepercayaan dan pengujian hipotesis dilakukan secara bersamaan. Itu sebabnya pada topik selang kepercayaan kami belum menjelaskan penggunaan

Program Minitab. Kemudian opsi satu lagi adalah *Alternative hypothesis* yang *default*nya adalah *Mean ≠ hypothesized mean* (Gambar 5.2.1.6).

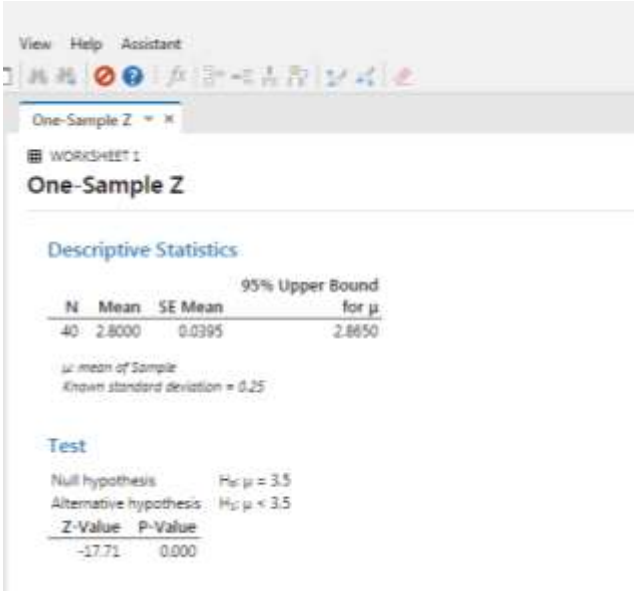
Gambar 5.2.1.6 Kotak dialog *One-Sample Z* : *Options* (1)

Klik *Mean ≠ hypothesized mean* dan ganti dengan *Mean < hypothesized mean* (Gambar 5.2.1.7).



Gambar 5.2.1.7 Kotak dialog *One-Sample Z* : *Options*

Lalu klik Ok sebanyak dua kali dan hasilnya diberikan pada Gambar berikut.



Gambar 5.2.1.8 *Output* Uji *One-Sample Z* Contoh 1

Pada Gambar terdapat 2 bagian *output* yaitu:

a. *Descriptive Statistics* (Statistik Deskriptif) yang di dalamnya terdapat 95% *Upper Bound for* μ yang merupakan batas atas selang kepercayaan 95%. Informasi ini menjelaskan bahwa nilai parameter yang akan diuji yaitu $\mu_0 = 3{,}5$ sudah melewati batas atas selang kepercayaan untuk rataan μ, yang artinya $\mu_0 = 3{,}5$ terletak di luar selang kepercayaan. Jika parameter yang akan diuji terletak di luar selang kepercayaan, sudah dipastikan H_0 ditolak. Ini merupakan salah satu cara melihat dengan cepat apakah H_0 ditolak atau tidak.
b. *Test* (Pengujian Hipotesis)

   Pada bagian ini terlihat hasil pengujian hipotesis. Kita perhatikan bagian Tabel 5.2.1.1 berikut.

Tabel 5.2.1.1 Hasil pengujian hipotesis Contoh 1

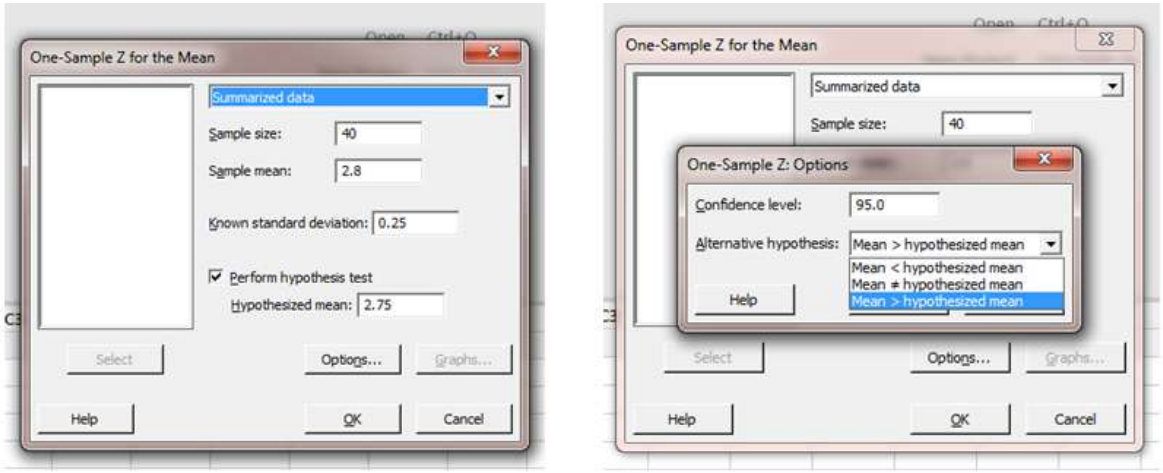
Z-*Value*	P-*Value*
–17,71	0,000

Z-*Value* adalah nilai statistik Z (z_{sampel}) dengan P-*value* (nilai-p) = 0,000.

Hasil dan kesimpulannya sama dengan cara yang dikerjakan sebelumnya secara manual.

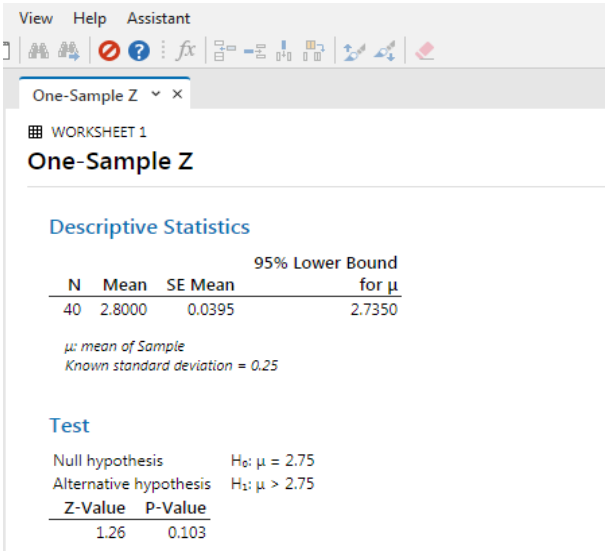
Contoh 2.

Sekarang kita ulang mengerjakan Contoh 1 , kita akan menguji H_0 : μ = 2,75 ; perhatikan nilai μ_0 = 2,75 ini terletak dalam selang kepercayaan 95% (2,72 < μ < 2,88). Akan kita lihat bahwa H_0 tidak dapat ditolak (diterima). Berikut ini pengujian hipotesis H_0 : μ = 2,75 yang dikerjakan dengan menggunakan Program Minitab. Langkahnya sama dengan pada Contoh 1, tetapi sekarang kita isikan data *Hypothesized mean* = 2,75 dan *Options Alternative Hypothesis* pilih *Mean > hypothesized mean*, karena rataan sampel lebih besar dari 2,75(Gambar 5.2.1.9).



Gambar 5.2.1.9 Uji hipotesis dengan alternatif *Mean > hypothesized mean*

Hasilnya diberikan pada Gambar 5.2.1.10 berikut.



Gambar 5.2.1.10 *Output* Uji *One-Sample Z* Contoh 2

Pada bagian *Descriptive Statistics* terlihat diberikan 95% *Lower Bound for* μ yang merupakan batas bawah selang kepercayaan 95% untuk μ. Informasi ini menjelaskan bahwa nilai parameter yang akan diuji yaitu $\mu_0 = 2{,}75$ lebih besar nilainya dari batas bawah selang kepercayaan 95% yang berarti terletak di dalam selang kepercayaan 95%. Itulah sebabnya H_0 diterima. Lain halnya pada Contoh 1, nilai parameter yang akan diuji yaitu $\mu_0 = 3{,}5$ lebih besar nilainya dari batas atas selang kepercayaan 95% yang berarti terletak di luar selang kepercayaan 95%.

Pada bagian *Test*, hasil pengujian hipotesis ditulis ulang dalam Tabel 5.2.1.2 berikut.

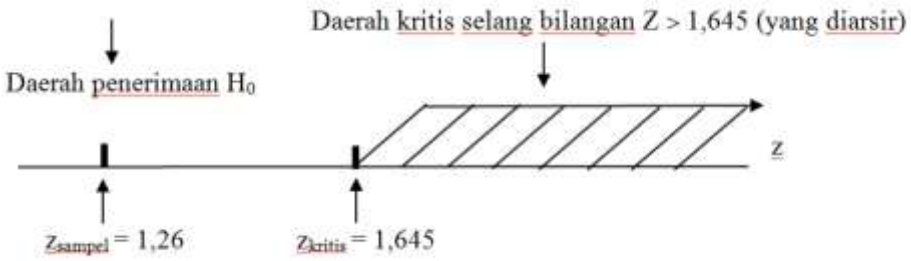
Tabel 5.2.1.2 Hasil pengujian hipotesis Contoh 2

Null hypothesis (H_0)	H_0: $\mu = 2{,}75$
Alternative hypothesis (H_1)	H_1: $\mu > 2{,}75$

Z-*Value*	P-*Value*
1,26	0,103

Dari Tabel 5.2.1.2 terlihat bahwa Z-*Value* = z_{sampel} = 1,26 dengan P-*value* (nilai-p) = 0,103.
Dalam Contoh 2 ini, berdasarkan relasi $H_1 : \mu > 2{,}75$, maka daerah kritis adalah,

$Z > z_\alpha$ yaitu $Z > z_{0,05} \rightarrow Z > 1{,}645$



Gambar 5.2.1.11 Daerah kritis Contoh 2

Dalam kasus ini: nilai-p = $P(Z > z_{sampel}) = P(Z > 1{,}26) = 0{,}103$. Maka kita dapat membuat kesimpulan berdasarkan dua kriteria:

a. Berdasarkan bahwa nilai z_{sampel} terletak pada daerah penerimaan H_0.
   Maka H_0 diterima bahwa kita masih bisa menerima rataan nilai Statistika adalah 2,75.
b. Berdasarkan nilai-p = 0,103 > α = 0,05 , maka H_0 diterima.

\

Contoh 3

Sekarang kita kerjakan kembali soal nomor 1, tetapi dengan hipotesis alternatif $\mu \neq \mu_0$.

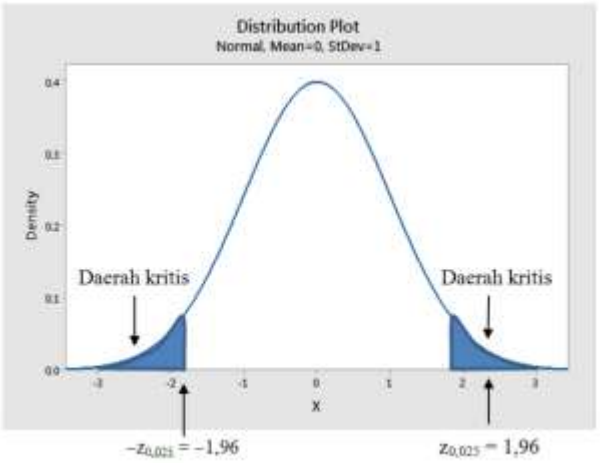
1. H_0 : $\mu = 3{,}5$
2. H_1 : $\mu \neq 3{,}5$ $\rightarrow$ Uji dwi-arah
3. $\alpha = 0{,}05$
4. Uji statistik yang sesuai : $Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$; karena n > 40, σ dapat diganti s = 0,25 .

   Dengan $\overline{X}$ adalah rata-rata sampel, yaitu : $\overline{X}$ = 2,8.

   Karena H_1 : $\mu \neq 3{,}5$ merupakan uji dwi-arah, maka daerah kritis, $Z < -z_{0,025}$ dan $Z > z_{0,025}$ yaitu,

   $Z < -1{,}96$ dan $Z > 1{,}96$

   seperti yang diilustrasikan dalam Gambar 5.2.1.12 berikut.



Gambar 5.2.1.12 Daerah kritis Contoh 3

5. Perhitungan nilai statistik dari sampel acak ukuran n (sama seperti sebelumnya):

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{2{,}8 - 3{,}5}{0{,}25 / \sqrt{40}} = -17{,}71 \rightarrow z_{sampel}$$

6. Perhatikan bahwa Z_{sampel} terletak di daerah kritis yang sebelah kiri yaitu,

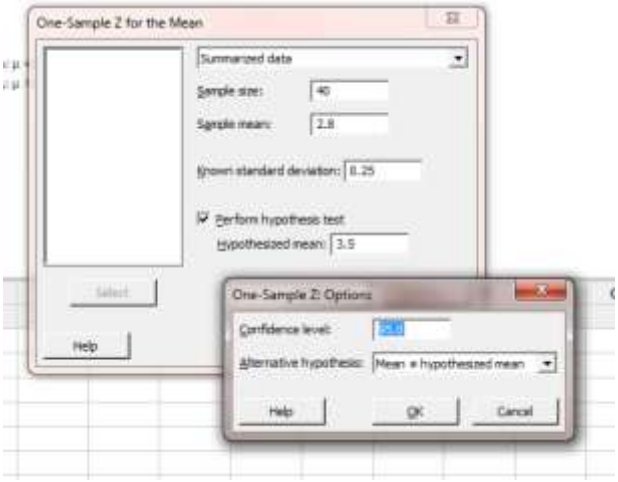
   $Z_{hitung} < Z_{kritis\ (kiri)}$ maka H_0 ditolak

   Kesimpulan : Rataan nilai Statistika mahasiswa tingkat sarjana adalah lebih kecil dari 3,5.

   Jadi kita menolak hipotesis dosen Statistika tersebut.

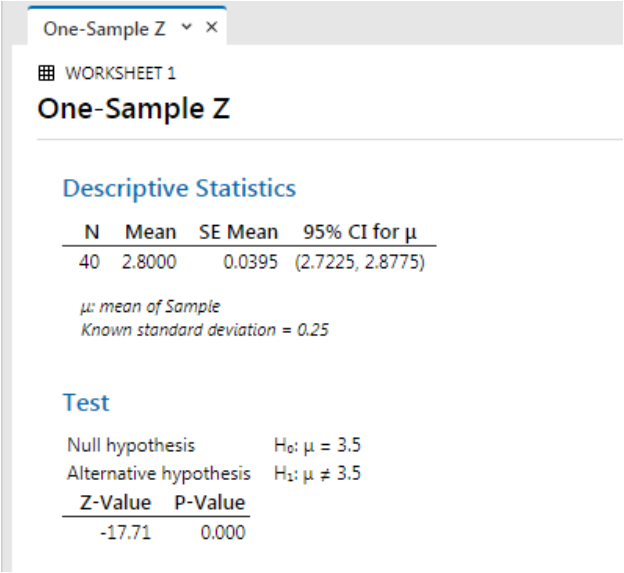
   Dalam kasus ini daerah kritis terletak di ujung kiri dan ujung kanan distribusi. Maka kita akan menolak H_0 apabila $Z_{hitung} < Z_{kritis\ (kiri)}$ atau $Z_{hitung} > Z_{kritis\ (kanan)}$

Berikut ini diberikan prosedur penentuan selang kepercayaan dan pengujian hipotesis Contoh 3 dengan menggunakan Program Minitab. Langkahnya sama seperti pada Contoh 1, tetapi pada *Alternative hypothesis* pilih *Mean ≠ Hypothesized mean* (Gambar 5.2.1.13).



Gambar 5.2.1.13 Uji hipotesis dengan alternatif *Mean ≠ Hypothesized mean*

Hasilnya diberikan pada Gambar 5.2.1.14 berikut.



Gambar 5.2.1.14 *Output* Uji *One-Sample Z* Contoh 3

Pada bagian *Descriptive Statistics* terlihat diberikan 95% *CI for* μ (selang kepercayaan 95% untuk μ) yaitu 2,72 – 2,88. Jadi jika kita menggunakan uji dwiarah, maka Program Minitab akan menampilkan selang kepercayaan 95%. Dengan ini diharapkan semakin memahami penggunaan menu yang ada pada Program Minitab dalam prosedur penentuan selang kepercayaan dan pengujian hipotesis yang dilakukan secara bersamaan.

Dalam Contoh 1 sampai Contoh 3 kita telah membahas tiga prosedur pengujian hipotesis dengan hipotesis alternatif H_1 ekaarah ($\theta < \theta_0$ dan $\theta > \theta_0$) dan dwiarah ($\theta \neq \theta_0$). Sekarang kita akan membahas bagaimana kita memilih hipotesis alternatif H_1 sebagai berikut:

a. $H_1 : \theta < \theta_0$ jika nilai penaksir $\hat{\theta}$ lebih kecil dari nilai parameter θ_0 yang akan diuji.

b. $H_1 : \theta > \theta_0$ jika nilai penaksir $\hat{\theta}$ lebih besar dari nilai parameter θ_0 yang akan diuji.

c. $H_1 : \theta \neq \theta_0$ dapat digunakan jika nilai penaksir $\hat{\theta}$ lebih kecil atau lebih besar dari nilai parameter θ_0 yang akan diuji. Pilihan c ini merupakan *default* dalam *Perform hyphotesis test* pada Program Minitab.

Sekarang kita akan melihat apa yang terjadi jika misalnya kita memilih $H_1 : \theta > \theta_0$ sementara nilai penaksir $\hat{\theta}$ lebih kecil dari nilai parameter θ_0 yang akan diuji. Kita akan menggunakan Contoh 1. Prosedur pengujian hipotesis sebagai berikut.

1. $H_0 : \mu = 3,5$
2. Sekarang kita coba menggunakan $H_1 : \mu > \mu_0$ maka $H_1 : \mu > 3,5$
3. $\alpha = 0,05$
4. Uji statistik yang sesuai masih sama seperti pada Contoh 1 yaitu,

   $$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$; n = 40 > 30, σ dapat diganti s = 0,25

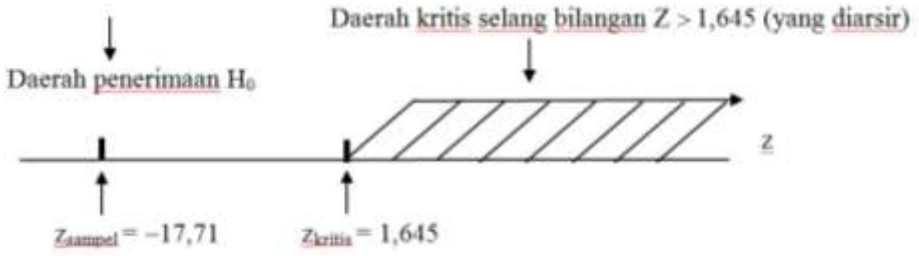
   Dengan $\overline{X}$ adalah rata-rata sampel, yaitu : $\overline{X}$ = 2,8.

   Dari Tabel nomor 1 kolom 4 dan 5, berdasarkan H_1 maka sekarang daerah kritis adalah ujung sebelah kanan yaitu $Z > z_{0,05}$ yaitu Z > 1,645.

5. Nilai statistik dari sampel acak ukuran n = 40 masih sama seperti pada Contoh 1 yaitu,

   $$z_{sampel} = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{2,8 - 3,5}{0,25 / \sqrt{40}} = -17,71$$

6. Kesimpulan dapat dilakukan sebagai berikut. Pada Gambar 5.2.1.15 terlihat bahwa nilai statistik Z (z_{sampel}) terletak dalam daerah penerimaan H_0, oleh karenanya H_0 diterima.



Gambar 5.2.1.15 z_{sampel} terletak dalam daerah penerimaan H_0

Dalam kasus ini kita harus menerima H_0 bahwa $\mu = 3,5$ dari pada menolak H_0 lalu menerima H_1 bahwa $\mu > 3,5$ yang nilainya semakin jauh dari nilai penaksir titknya. Pemilihan hipotesis alternatif H_1 seperti ini dapat menyebabkan kesimpulan yang tidak tepat. Itu sebabnya dalam melakukan inferensi seharusnyalah sampel yang diambil benar-benar merepresentasikan populasi, dan itu harus didukung oleh prosedur pengujian hipotesis yang benar.

5.2.2. Uji hipotesis rata-rata suatu populasi ($n < 30$)

Sekarang kita lanjutkan ke prosedur pengujian hipotesis rata-rata populasi untuk ukuran sampel kecil. Soal yang digunakan adalah soal dalam pembahasan selang kepercayaan untuk rata-rata populasi ukuran sampel kecil yaitu Contoh 4.5.2 pada halaman 72.

Contoh 4.

Di sebuah depot air minum isi ulang dilakukan percobaan untuk menaksir rata-rata volume galon air (dalam liter). Percobaan dilakukan terhadap 7 galon air dan volumenya adalah 19,1 19.2 19.4 19.1 18.9 19.2 19.0 liter. Ujilah hipotesis bahwa rata-rata volume semua galon air pada depot tersebut adalah 18 liter dengan hipotesis alternatif bahwa rata-rata sesungguhnya lebih besar dari 18. Gunakan taraf signifikansi 0,05. Diasumsikan distribusi populasi hampir normal.

Penyelesaian

Dari soal selang kepercayaan sudah diperoleh bahwa rata-rata dan standar deviasi sampel di atas:

$\bar{x} = 19,129$ dan $s = 0,160$

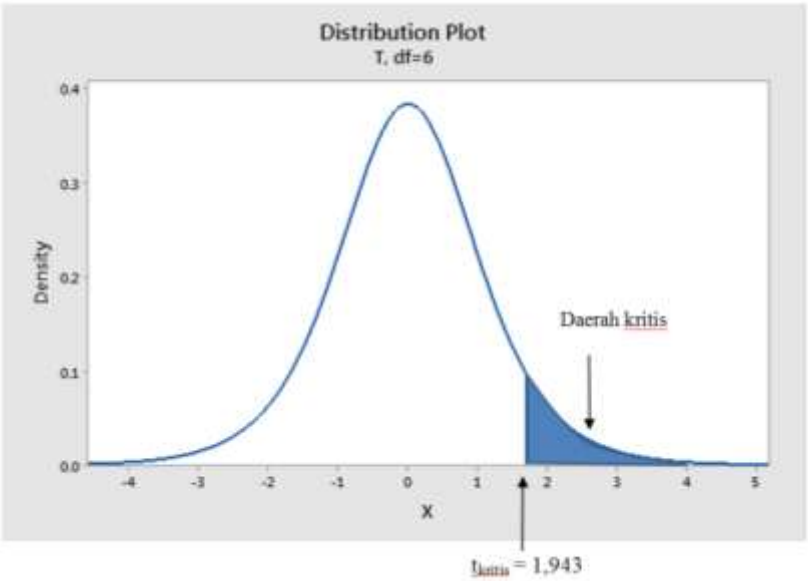
Langkah pengujian hipotesis.

1. $H_0 : \mu = 18$
2. $H_1 : \mu > 18$
3. $\alpha = 0,05$

4. Dari Tabel 5.2.1 gunakan uji statistik nomor 2 yaitu,

$$T = \frac{\overline{X} - \mu_0}{s/\sqrt{n}}$$; n = 7 maka $\nu = n - 1 = 7 - 1 = 6$

Dari Tabel nomor 2 kolom 4 dan 5, dan berdasarkan H_1 maka daerah kritis adalah ujung sebelah kanan yaitu $T > t_\alpha$ atau $T > t_{0,05,\nu=6}$ di mana $t_{0,05,\ \nu\ =\ 6} = 1{,}943$. Jadi daerah kritis adalah $T > 1{,}943$ seperti yang diilustrasikan pada Gambar 5.2.2.1 berikut.



Gambar 5.2.2.1 Daerah kritis Contoh 4

5. Perhitung nilai statistik dari sampel acak ukuran n = 7 :

$$t_{sampel} = \frac{19{,}129 - 18}{0{,}16/\sqrt{7}} = 18{,}62 \rightarrow \text{terletak pada daerah kritis}$$

6. Perhatikan bahwa : $t_{sampel} > t_{kritis} \Rightarrow$ Tolak H_0

Kesimpulan : Kita menolak hipotesis bahwa rata-rata volume galon adalah 18 liter dan kita terima bahwa rata-rata isi volume galon pada depot air isi ulang tersebut lebih besar dari 18 liter.

Langkah 1.

Input data pada *Worksheet* (lembar kerja). Beri nama Volume Galon. Biasanya dalam Program minitab penulisan angka desimal menggunakan tanda titik seperti pada Gambar 5.2.2.2. Jika

memang benar menggunakan tanda titik, maka pada baris "C1 C2 C3 dan seterusnya" di kolom yang sesuai tidak ada huruf T, seperti kolom data C1 pada Gambar 5.2.2.2 berikut.

	C1	C2	C3
	Volume Galon		
1	19.1		
2	19.2		
3	19.4		
4	19.1		
5	18.9		
6	19.2		
7	19.0		

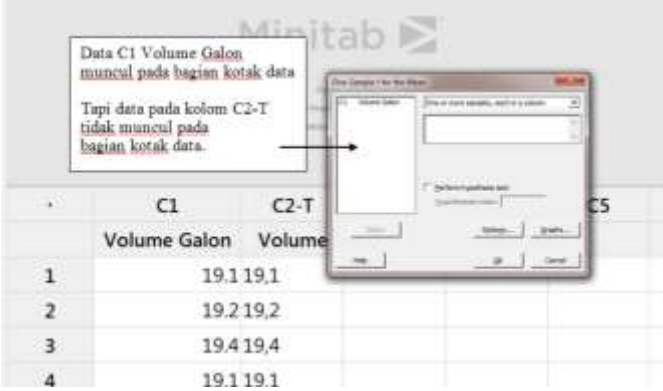
Gambar 5.2.2.2 Input data pada *worksheet* untuk Contoh 4

Jika kita menggunakan tanda koma padahal seharusnya tanda titik seperti pada Gambar 5.2.2.3 kolom C2, maka pada baris "C1 C2 C3 dan seterusnya" di kolom yang sesuai (seperti C2 pada Gambar 5.2.2.3) akan menjadi C2-T. T ini bisa berarti *Time* atau *Text* yang menyatakan data kita bukan data numerik atau angka.

	C1	C2-T	C3
	Volume Galon	Volume	
1	19.1	19,1	
2	19.2	19,2	
3	19.4	19,4	
4	19.1	19,1	
5	18.9	18,9	
6	19.2	19,2	
7	19.0	19,0	

Gambar 5.2.2.3 Input data desimal yang salah

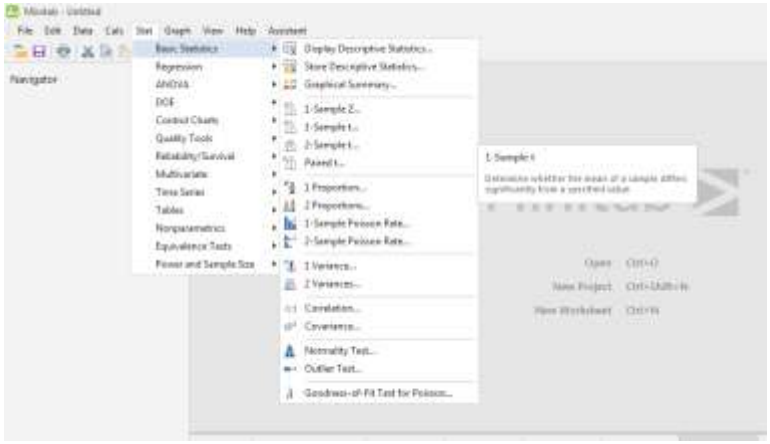
Jadi harus diperhatikan versi Program Minitab yang kita gunakan, apakah menggunakan tanda titik atau tanda koma untuk angka desimal. Dan harus diingat bahwa Program Minitab mengolah data angka numerik. Pembaca bisa melihatnya ketika hendak memproses data, maka data "*Time*" atau data "*Text*" tidak muncul di kotak data (Gambar 5.2.2.4).



Gambar 5.2.2.4 Data *Time* atau *Text* tidak muncul pada kotak data

Langkah 2.

Dari ikon *Stat*, pilih *Basic Statistics*, kemudian klik *1-Sample t* (Gambar 5.2.2.5).

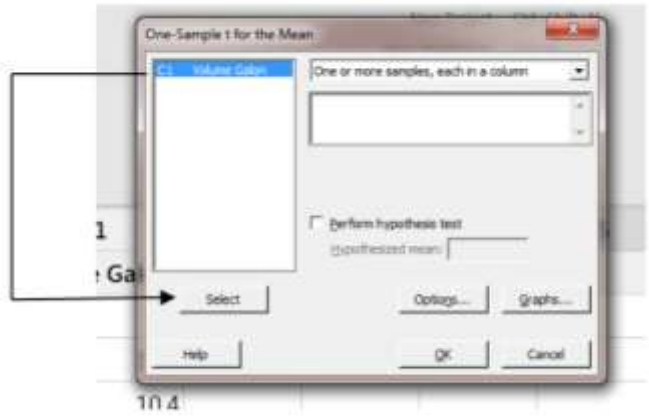


Gambar 5.2.2.5 Menu 1-*Sample t* Program Minitab

Akan muncul kotak dialog *One-Sample* t *for the Mean* seperti pada Gambar 5.2.2.6, kemudian pilih *One or more samples, each in a column*. Pilihan ini karena data yang diberikan merupakan data mentah yang diinput pada kolom C1.

Gambar 5.2.2.6 Kotak dialog *One-Sample* t *for the Mean*

Langkah 2. Setelah diklik *One or more samples, each in a column* klik C1 Volume Galon (ada di kotak sebelah kiri atas, kita sebut kotak data) lalu klik *Select* (Gambar 5.2.2.7).



Gambar 5.2.2.7 Langkah 2 Uji *One-Sample* t *for the Mean*

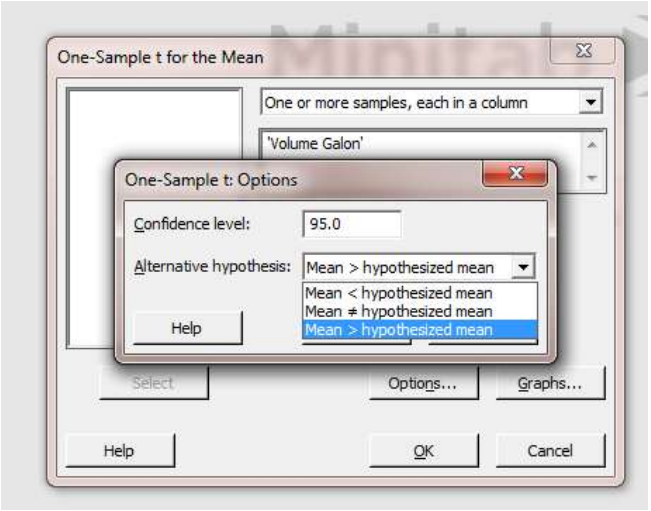
Hasilnya diberikan pada Gambar 5.2.2.8 berikut.

Gambar 5.2.2.8 Hasil Langkah 2 Uji *One-Sample* t *for the Mean*

Kemudian centang *Perform hyphotesis test* dan isikan data *Hypothesized mean* (nilai rataan di bawah H_0) yaitu 18 (Gambar 5.2.2.9).

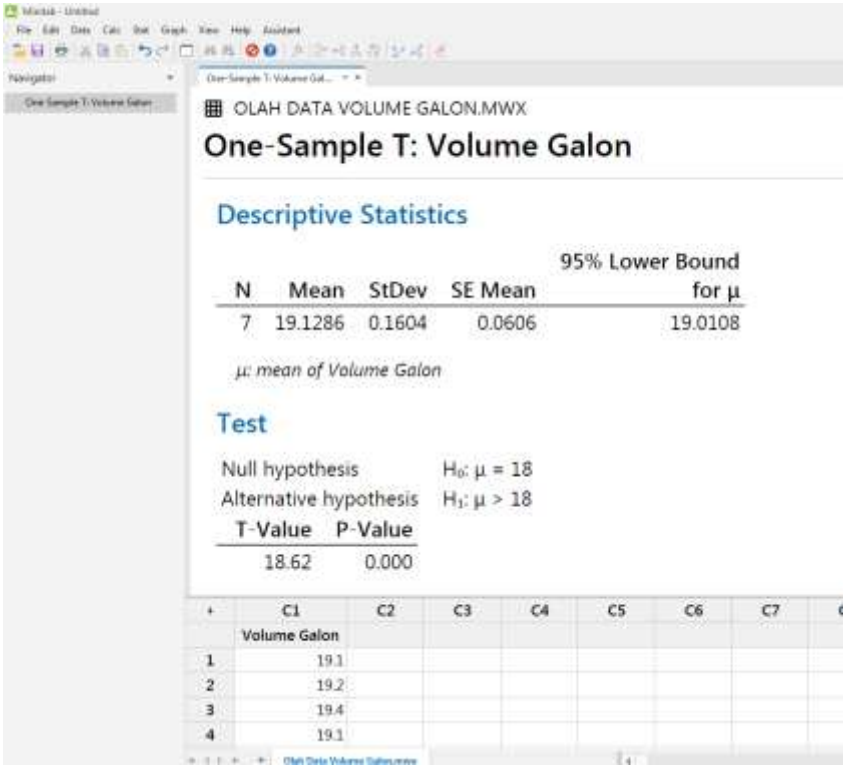
Gambar 5.2.2.9 *Perform hyphotesis test* Contoh 4

Lalu klik *Options*. Seperti pada Contoh 1 uji *One-Sample Z for the Mean*, *default Confidence level* (selang kepercayaan) adalah 95% sesuai taraf signifikansi α = 5%. Pilih *Alternative hypothesis Mean > hypothesized mean* (Gambar 5.2.2.10).



Gambar 5.2.2.10 Kotak dialog *One-Sample* t: *Options*

Kemudian klik Ok dua kali, hasilnya diberikan pada Gambar 5.2.2.11 berikut.



Gambar 5.2.2.11 *Output* Uji *One-Sample t* Contoh 4

Dari Gambar di atas kita peroleh:

a. Di dalam bagian *Descriptive Statistics* (Statistik Deskriptif) terdapat 95% *Lower Bound for* μ yang merupakan batas bawah selang kepercayaan 95%. Informasi ini menjelaskan bahwa nilai parameter yang akan diuji yaitu $\mu_0 = 18$ lebih kecil dari batas bawah selang kepercayaan untuk rataan μ, yang artinya $\mu_0 = 18$ terletak di luar selang kepercayaan. Oleh karenanya H_0 ditolak.

b. *Test* (Pengujian Hipotesis)

Pada bagian ini terlihat hasil pengujian hipotesis. Kita perhatikan bagian Tabel 5.2.2 berikut.

Tabel 5.2.2 Hasil pengujian hipotesis Contoh 4

T-*Value*	P-*Value*
18,62	0.000

T-*Value* adalah nilai statistik T (t_{sampel}) dengan P-*value* (nilai-p) = 0,000.

Hasil dan kesimpulannya sama dengan cara yang dikerjakan sebelumnya secara manual. Kita menolak hipotesis bahwa rata-rata volume galon adalah 18 liter dan kita terima bahwa rata-rata isi volume galon pada depot air isi ulang tersebut lebih besar dari 18 liter.

5.2.3. Uji hipotesis selisih rata-rata dua populasi bebas ($n_1 \geq 30$ dan $n_2 \geq 30$)

Contoh 5.

Pada masa pandemi dilaksanakan kuliah Statistika secara daring dengan jumlah peserta yang besar terdiri dari 80 mahasiswa wanita dan 60 mahasiswa pria. Nilai rata-rata Ujian Tengah Semester (UTS) mahasiswa wanita adalah 85 dengan standar deviasi 6, sedangkan nilai rata-rata UTS mahasiswa pria adalah 78 dengan standar deviasi 7. Dari sini kita melihat bahwa taksiran titik untuk $\mu_1 - \mu_2$: $\bar{x} - \bar{y} = 85 - 78 = 7$. Sekarang kita akan menyusun prosedur pengujian hipotesis untuk selisih $\mu_1 - \mu_2$ di mana μ_1 menyatakan rata-rata nilai semua mahasiswa wanita dan μ_2 menyatakan rata-rata nilai semua mahasiswa pria yang mengikuti ujian matakuliah ini. Pada Contoh 4.5.3 halaman 74 selang kepercayaan 96% adalah $4{,}69 < \mu_1 - \mu_2 < 9{,}31$. Ujilah hipotesis nol bahwa selisih $\mu_1 - \mu_2 = d_0 = 10$ dengan hipotesis alternatif $\mu_1 - \mu_2 \neq 10$. Perhatikan bahwa nilai selisih ini terletak di luar selang kepercayaan 96%. Kita akan menggunakan taraf keberartian $\alpha = 0{,}04$ karena selang kepercayaan dalam contoh ini adalah 96%.

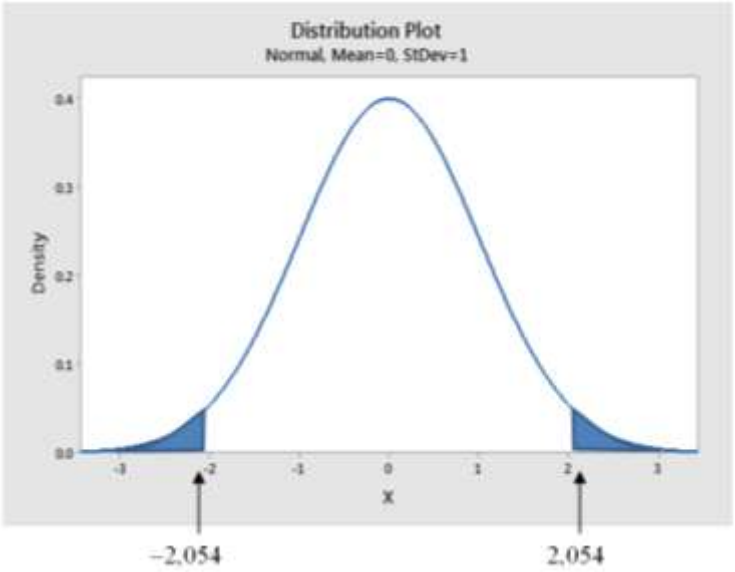
Penyelesaian

Langkah pengujian hipotesis.

1. $H_0 : \mu_1 - \mu_2 = d_0 = 10$
2. $H_1 : \mu_1 - \mu_2 \neq 10$
3. $\alpha = 0{,}04$
4. Dari Tabel 5.2.1 gunakan uji statistik nomor 3 yaitu,

$$Z = \frac{(\overline{X} - \overline{Y}) - d_0}{\sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}}$$

Berdasarkan H_1 maka daerah kritis adalah ujung kiri dan ujung kanan yaitu $Z < -z_{\alpha/2}$ dan $Z > z_{\alpha/2}$. Karena Selang kepercayaan 96%, $\Rightarrow 1 - \alpha = 96\% \Rightarrow \alpha = 4\% = 0{,}04 \Rightarrow \alpha/2 = 0{,}02$ dan dengan menggunakan Program Minitab $z_{0{,}02} = 2{,}054$. Maka daerah kritis adalah $Z < -2{,}054$ dan $Z > 2{,}054$ seperti yang diilustrasikan pada Gambar 5.2.3.1 erikut.



Gambar 5.2.3.1 Daerah kritis Contoh 5

5. Taksiran titik untuk $\mu_1 - \mu_2$: $\overline{x} - \overline{y} = 85 - 78 = 7$.

Karena n_1 dan n_2 keduanya besar, σ_1 dapat diganti $s_1 = 6$ dan σ_2 dengan $s_2 = 7$.

Maka nilai statistik,

$$z_{sampel} = \frac{7 - 10}{\sqrt{(36/80) + (49/60)}} = -2{,}67$$

6. Kesimpulan : Z_{sampel} terletak pada daerah kritis di sebelah kiri, maka H_0 ditolak dan kita menerima H_1 yaitu selisih rata-rata nilai semua mahasiswa wanita dengan rata-rata nilai semua mahasiswa pria yang mengikuti ujian matakuliah Statistika ini tidak sama dengan 10.

Catatan.

Prosedur penentuan selang kepercayaan dan pengujian hipotesis selisih rata-rata dua populasi ukuran sampel besar tidak tersedia pada Program Minitab sehingga dikerjakan secara manual.

5.2.4. Uji hipotesis selisih rata-rata dua populasi bebas ($n_1 < 30$ dan $n_2 < 30$**)**

a. Variansi sama tetapi tidak diketahui.

Contoh 6.

Dalam sebuah pelaksanaan perkuliahan secara daring dengan jumlah peserta yang kecil, hendak dibandingkan rata-rata nilai UTS matakuliah Matematika I dengan rata-rata nilai UTS matakuliah Statistika. Nilai rata-rata UTS 18 mahasiswa peserta Matematika I adalah 82 dengan standar deviasi 5, sedangkan nilai rata-rata UTS 15 mahasiswa peserta Statistika adalah 75 dengan standar deviasi 4. Perhatikan bahwa taksiran titik untuk $\mu_1 - \mu_2$ adalah,

$$\bar{x}_1 - \bar{x}_2 = 82 - 75 = 7$$

di mana μ_1 dan μ_2 masing-masing menyatakan rata-rata populasi semua peserta yang mungkin untuk kedua matakuliah. Dari Contoh 4.5.4 halaman 77 selang kepercayaan 96% untuk $\mu_1 - \mu_2$ adalah $3{,}57 < \mu_1 - \mu_2 < 10{,}43$. Sekarang kita akan menyusun prosedur pengujian hipotesis untuk selisih $\mu_1 - \mu_2$. Ujilah hipotesis yang berpendapat bahwa selisih $\mu_1 - \mu_2$ adalah 0 yang berarti $\mu_1 = \mu_2$ dengan hipotesis alternatif $\mu_1 - \mu_2 \neq 0$ atau $\mu_1 \neq \mu_2$. Gunakan $\alpha = 4\% = 0{,}04$. Diasumsikan kedua populasi berdistribusi hampir normal dengan variansi yang sama tetapi tidak diketahui.

Penyelesaian

Dari Contoh 4.5.7 halaman 77 diperoleh :

$s_p = 4{,}58$ dan dengan $\alpha = 4\% = 0{,}04 \Rightarrow \alpha/2 = 0{,}02 \Rightarrow t_{0{,}02\,,\,v=31} = 2{,}14$.

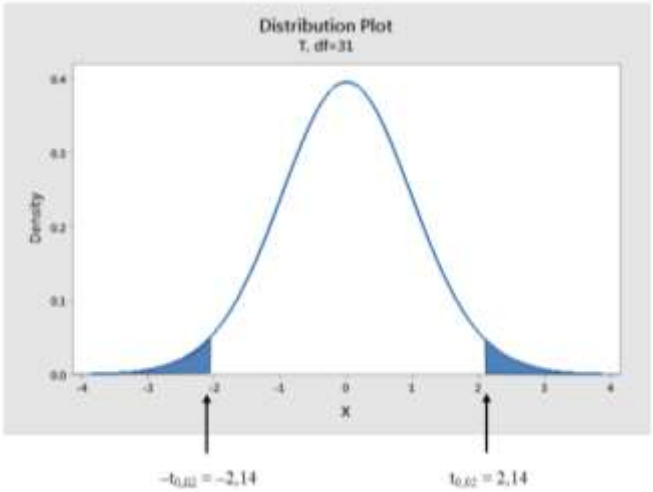
Langkah pengujian hipotesis.

1. $H_0 : \mu_1 - \mu_2 = 0$
2. $H_1 : \mu_1 - \mu_2 \neq 0$
3. $\alpha = 0{,}04$

4. Dari Tabel 5.2.1 gunakan uji statistik nomor 4 yaitu,

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{S_p\sqrt{(1/n_1) + (1/n_2)}} \text{ , } \nu = 18 + 15 - 2 = 31$$

Berdasarkan H_1 maka daerah kritis adalah ujung kiri dan ujung kanan yaitu $T < -z_{\alpha/2}$ dan $T > t_{\alpha/2}$. Dengan $t_{0,02;\ \nu=31} = 2,14$, maka daerah kritis adalah $T < -2,14$ dan $T > 2,14$ seperti yang diilustrasikan pada Gambar 5.2.4.1 berikut.



Gambar 5.2.4.1 Daerah kritis Contoh 6

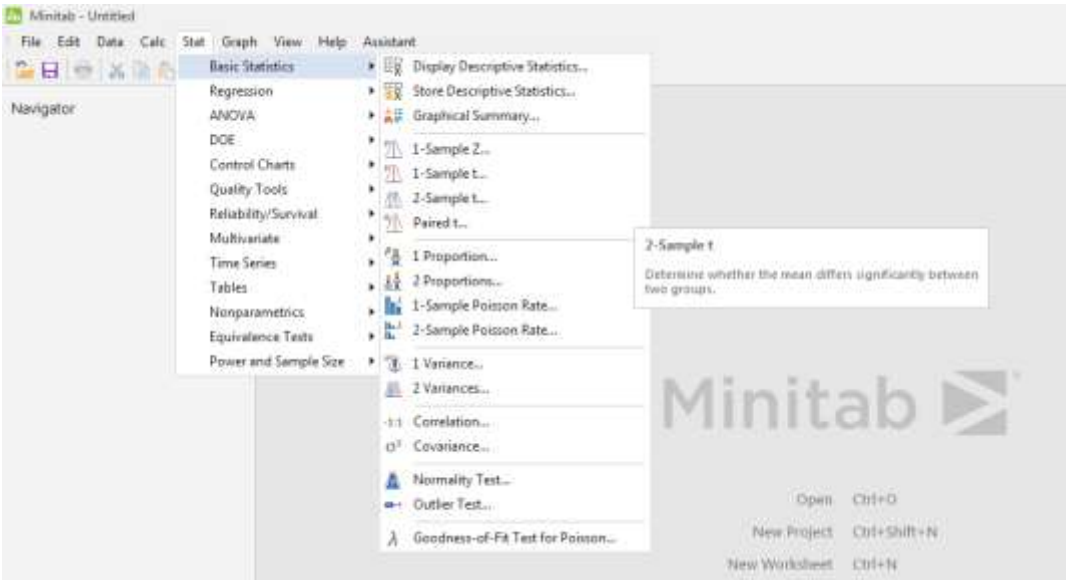
5. Nilai statistik,

$$t_{sampel} = \frac{7 - 0}{4,58\sqrt{(1/18) + (1/15)}} = 4,375$$

Kesimpulan : t_{sampel} terletak pada daerah kritis di sebelah kanan, maka H_0 ditolak dan kita menerima H_1 yaitu selisih $\mu_1 - \mu_2 \neq 0$. Hal ini bermakna bahwa selisih rata-rata nilai UTS matakuliah Matematika I dengan rata-rata nilai UTS matakuliah Statistika tidak sama dengan nol, atau rata-rata nilai UTS matakuliah Matematika I tidak sama dengan rata-rata nilai UTS matakuliah Statistika.

Berikut langkah uji hipotesis dengan menggunakan Program Minitab.

Langkah 1. Pilih ikon *Stat*, *Basic Statistics* dan 2-*Sample* t (Gambar 5.2.4.2).



Gambar 5.2.4.2 Menu 2-*Sample* t

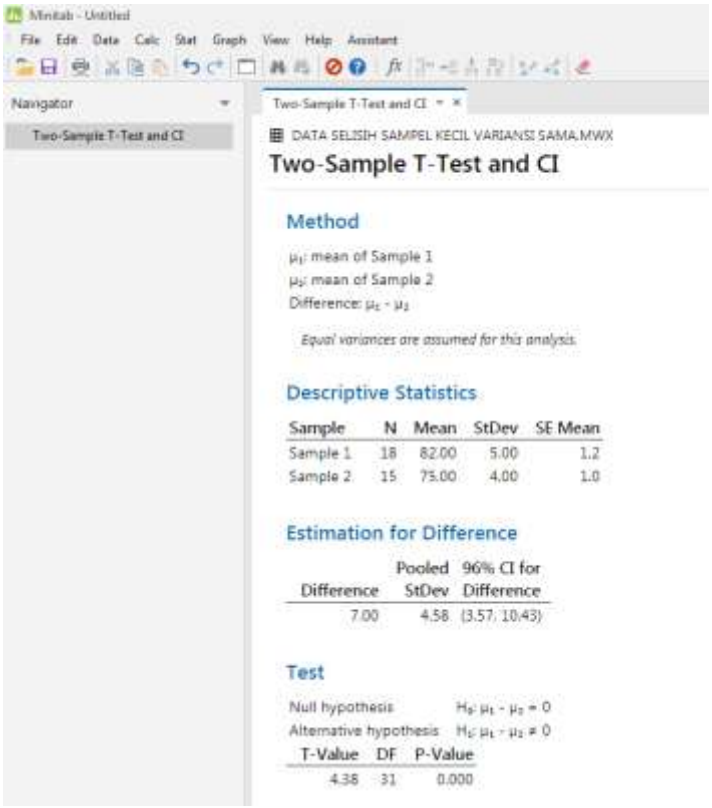
Langkah 2. Akan muncul kotak *Two-Sample* t *for the Mean*. Pilih *Summarized data* karena data yang kita miliki adalah ringkasan data. Kemudian isikan ringkasan data sesuai soal di atas, seperti yang diperlihatkan pada Gambar 5.2.4.3 berikut. Selanjutnya klik ikon *Options*.

Gambar 5.2.4.3 Kotak dialog *Two-Sample* t *for the Mean*

Langkah 3. Setelah ikon *Options* diklik akan muncul kotak dialog *Two-Sample* t: *Options*. Ganti *Confidence level* 96. *Hypothesized difference* dan *Alternative hypothesis* sudah sesuai dengan *default*. Centang *Assume equal variances*. Lalu klik Ok sebanyak dua kali (Gambar 5.2.4.4).

Gambar 5.2.4.4 Kotak dialog *Two-Sample* t: *Options*

Hasilnya diberikan pada Gambar 5.2.4.5 berikut.



Gambar 5.2.4.5 Hasil pengujian hipotesis Contoh 6

Dari Gambar diperoleh:

a. Selang kepercayaan 96% untuk selisih rata-rata nilai UTS matakuliah Matematika I dengan rata-rata nilai UTS matakuliah Statistika adalah $3{,}57 < \mu_1 - \mu_2 < 10{,}43$.

   Selang kepercayaan ini tidak memuat nilai selisih rata-rata yang dihipotesiskan yaitu,

   $H_0 : \mu_1 - \mu_2 = 0$

   Dari selang kepercayaan ini kita sudah bisa menyimpulkan bahwa H_0 ditolak.

b. T-*value* = 4,38 dengan P-*value* = 0,000 < α = 0,04.

   Maka H_0 ditolak bahwa selisih rata-rata nilai UTS matakuliah Matematika I dengan rata-rata nilai UTS matakuliah Statistika tidak sama dengan nol. Dengan perkataan lain rata-rata nilai UTS matakuliah Matematika I tidak sama dengan rata-rata nilai UTS matakuliah Statistika.

b. Variansi tidak sama dan tidak diketahui.

Contoh 7.

Hendak dibandingkan rata-rata berat badan bayi laki-laki dengan rata-rata berat badan bayi perempuan ketika lahir. Rata-rata berat badan dari 15 bayi laki-laki adalah 4 kg dengan standar deviasi 1kg, sedangkan rata-rata berat badan dari 10 bayi perempuan adalah 2,5 kg dengan standar deviasi 0,5 kg. Ujilah hipotesis bahwa selisih rata-rata $\mu_1 - \mu_2$ adalah 2,5 kg, apabila μ_1 adalah rata-rata berat badan populasi bayi laki-laki dan μ_2 adalah rata-rata berat badan populasi bayi perempuan. Pilih hipotesis alternatif $H_1 : \mu_1 - \mu_2 \neq 2{,}5$. Gunakan $\alpha = 0{,}05$. Anggap bahwa pengamatan berasal dari populasi normal dengan variansi yang berbeda dan tidak diketahui.

Penyelesaian

Langkah pengujian hipotesis.

1. $H_0 : \mu_1 - \mu_2 = 2{,}5$
2. $H_1 : \mu_1 - \mu_2 \neq 2{,}5$
3. $\alpha = 0{,}05$
4. Dari Tabel 5.2.1 gunakan uji statistik nomor 5 yaitu,

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{(S_1^2 / n_1) + (S_2^2 / n_2)}}$$

Pada Contoh 4.5.5 halaman 78 derajat bebas adalah,

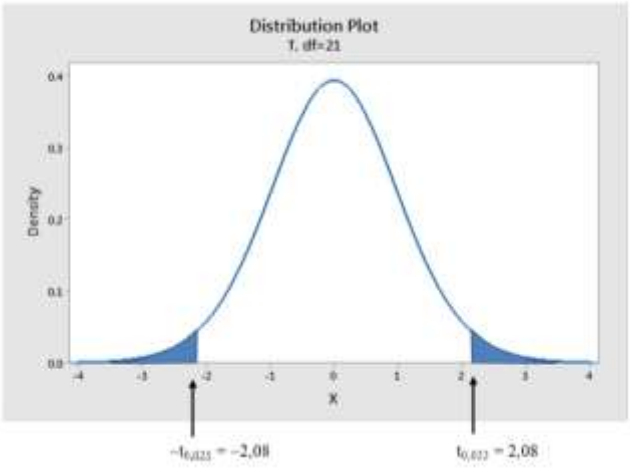
$$\nu = \frac{[(1^2/15)+(0{,}5^2/10)]^2}{[(1^2/15)^2/(15-1)]+[(0{,}5^2/10)^2/(10-1)]} \approx 21$$

Berdasarkan H_1 maka daerah kritis adalah ujung kiri dan ujung kanan yaitu $T < -t_{\alpha/2}$ dan $T > t_{\alpha/2}$. $\alpha = 0{,}05 \Rightarrow \alpha/2 = 0{,}025$ dan Dengan menggunakan Program Minitab diperoleh $t_{0{,}025}$ dengan $\nu = 21$: $t_{0{,}025\ ,\ \nu = 21} = 2{,}08$ (Tabel 5.2.4).

Tabel 5.2.4 Nilai $t_{0{,}025\ ,\ \nu = 21}$

$P(X \leq x)$	x
0.975	2.07961

Maka daerah kritis adalah $T < -2{,}08$ dan $T > 2{,}08$ seperti yang diilustrasikan pada Gambar 5.2.4.6 berikut.



Gambar 5.2.4.6. Daerah kritis Contoh 7

5. Taksiran titik untuk $\mu_1 - \mu_2$: $\bar{x}_1 - \bar{x}_2 = 4 - 2{,}5 = 1{,}5$

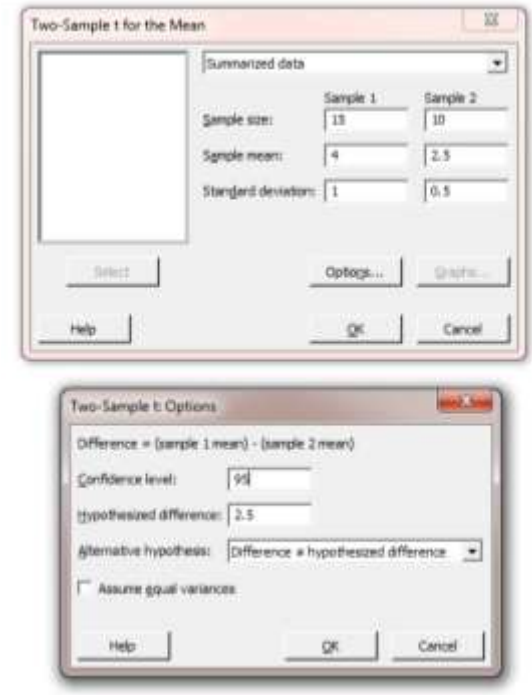
   Maka nilai statistik,

$$t_{sampel} = \frac{1{,}5 - 2{,}5}{\sqrt{(1^2/15)+(0{,}5^2/10)}} = -3{,}3$$

   Kesimpulan : t_{sampel} terletak pada daerah kritis di sebelah kiri, maka H_0 ditolak dan kita menerima H_1 yaitu selisih $\mu_1 - \mu_2 \neq 2{,}5$ bahwa selisih rata-rata berat badan populasi bayi laki-laki dengan rata-rata berat badan populasi bayi perempuan tidak sama dengan 2,5.

Langkah pengujian hipotesis dengan menggunakan Program Minitab untuk pengujian selisih rata-rata dua populasi ukuran sampel kecil variansi tidak sama dan tidak diketahui, sama dengan

langkah pengujian selisih rata-rata dua populasi ukuran sampel kecil variansi sama tetapi tidak diketahui. Pilih ikon *Stat*, *Basic Statistics* dan 2-*Sample* t. Setelah muncul kotak *Two-Sample* t *for the Mean*, pilih *Summarized data* karena data yang kita miliki adalah ringkasan data. Kemudian isikan ringkasan data sesuai soal di atasSelanjutnya klik ikon *Options*. Setelah ikon *Options* diklik akan muncul kotak *Two-Sample* t: *Options*. Isikan *Confidence level* 95. *Hypothesized difference* = 2,5 dan *Alternative hypothesis* sudah sesuai dengan *default*. Matikan centang *Assume equal variances*. Lalu klik Ok sebanyak dua kali (Gambar 5.2.7).



Gambar 5.2.7 Kotak dialog *Two-Sample* t *for the Mean*

Hasilnya diberikan pada Gambar 5.2.8 berikut.

Estimation for Difference

Difference	95% CI for Difference
1.500	(0.870, 2.130)

Test

Null hypothesis	$H_0: \mu_1 - \mu_2 = 2.5$
Alternative hypothesis	$H_1: \mu_1 - \mu_2 \neq 2.5$

T-Value	DF	P-Value
-3.30	21	0.003

Gambar 5.2.8 Hasil pengujian hipotesis Contoh 7

Dari Gambar diperoleh:

a. Selang kepercayaan 95% untuk selisih rata-rata berat badan populasi bayi laki-laki dengan rata-rata berat badan populasi bayi perempuan adalah 0,87 – 2,130.

   Selang kepercayaan ini tidak memuat nilai selisih rata-rata yang dihipotesiskan yaitu,

   $H_0 : \mu_1 - \mu_2 = 2{,}5$

   Dari selang kepercayaan ini kita sudah bisa menyimpulkan bahwa H_0 ditolak.

b. T-*value* = –3,3 dengan P-*value* = 0,003 < α = 0,05.

   Maka H_0 ditolak bahwa selisih rata-rata berat badan populasi bayi laki-laki dengan rata-rata berat badan populasi bayi perempuan tidak sama dengan 2,5.

5.2.5. Uji hipotesis selisih rata-rata data berpasangan

Contoh 8. (Contoh 4.5.6 halaman 80)

Duapuluh empat mahasiswa Teknik Mesin ITI dibagi menjadi 12 pasangan, tiap orang dalam pasangan mempunyai tingkat inteligensi yang hampir sama. Seorang dari tiap pasangan dipilih secara acak dan dimasukkan ke dalam kelompok yang diberi kuliah Matematika Teknik secara daring. Anggota lainnya masuk kelompok yang mengikuti kuliah secara luring. Pada akhir semester tiap kelompok diberi ujian yang sama dan nilai mereka diberikan pada Tabel berikut.

Tabel 5.2.5 Data pengamatan berpasangan (Contoh 8)

Pasangan	Daring	Luring	d
1	75	82	–7
2	64	59	5
3	82	87	–5
4	60	70	–10
5	85	90	–5
6	75	80	–5
7	82	90	–8
8	61	65	–4
9	79	87	–8
10	85	88	–3
11	79	86	–7
12	88	92	–4

Ujilah hipotesis bahwa rata-rata selisih nilai ujian kedua cara belajar adalah nol, dengan perkataan lain rata-rata nilai ujian kedua cara belajar adalah sama. Gunakan α = 2%.

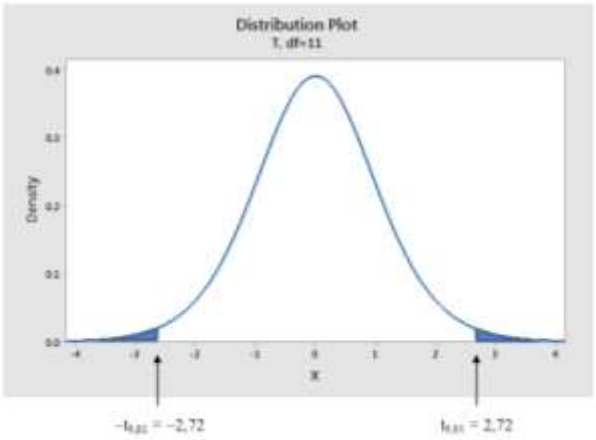
Penyelesaian

Langkah pengujian hipotesis.

1. $H_0 : \mu_D = 0$
2. $H_1 : \mu_D \neq 0$
3. $\alpha = 0{,}02$
4. Dari Tabel 5.2.1 gunakan uji statistik nomor 6 yaitu,

$$T = \frac{\overline{D} - d_0}{S_D / \sqrt{n}} \ , \ \nu = n - 1 = 12 - 1 = 11$$

Berdasarkan H_1 maka daerah kritis adalah ujung kiri dan ujung kanan yaitu $T < -z_{\alpha/2}$ dan $T > t_{\alpha/2}$. $\alpha = 0{,}02 \Rightarrow \alpha/2 = 0{,}01$ dan dengan menggunakan Program Minitab $t_{0{,}01\ ;\nu=11} = 2{,}72$. Maka daerah kritis adalah $T < -2{,}72$ dan $T > 2{,}72$ seperti yang diilustrasikan pada Gambar 5.2.5.1 berikut.



Gambar 5.2.5.1 Daerah kritis Contoh 8

5. Karena pengamatan dipasangkan, maka : $\mu_1 - \mu_2 = \mu_D$, dengan taksiran titik diberikan oleh

$$\bar{d} = \frac{-7 + 5 + ... + (-7) + (-4)}{12} = -5{,}08$$

Dan,

$$s_d^2 = \frac{n\sum_{i=1}^{12} d_i^2 - \left(\sum_{i=1}^{12} d_i \right)^2}{n(n-1)} = 14{,}27$$

$$\Rightarrow s_d = \sqrt{14{,}27} = 3{,}78$$

Maka nilai statistik,

$$t_{sampel} = \frac{-5{,}08 - 0}{3{,}78 / \sqrt{12}} = -4{,}66$$

6. Kesimpulan : t_{sampel} terletak pada daerah kritis di sebelah kiri, maka H_0 ditolak dan kita menerima H_1 yaitu selisih $\mu_D \neq 0$.

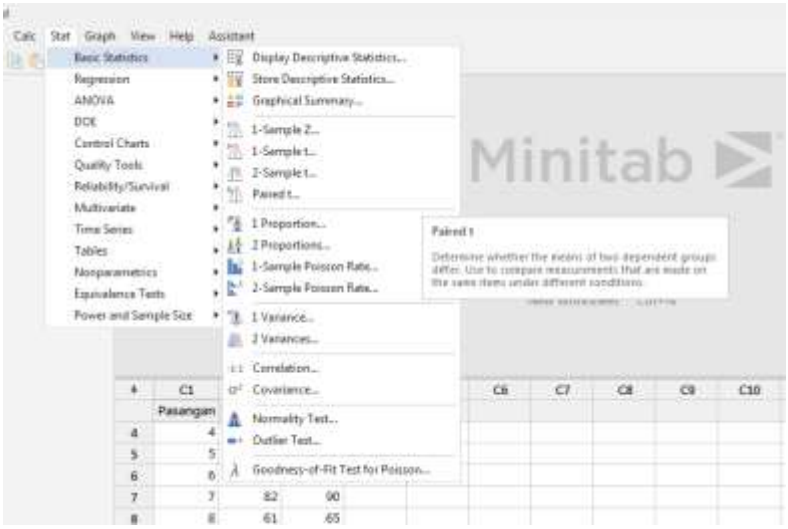
Berikut langkah pengujian hipotesis dengan menggunakan Program Minitab.

Langkah 1. Input data pada lembar kerja (*worksheet*) seperti yang diperlihatkan pada Gambar 5.2.5.2 berikut.

	C1	C2	C3	C4
	Pasangan	Daring	Luring	
4	4	60	70	
5	5	85	90	
6	6	75	80	
7	7	82	90	
8	8	61	65	
9	9	79	87	
10	10	85	88	
11	11	79	86	
12	12	88	92	
13				

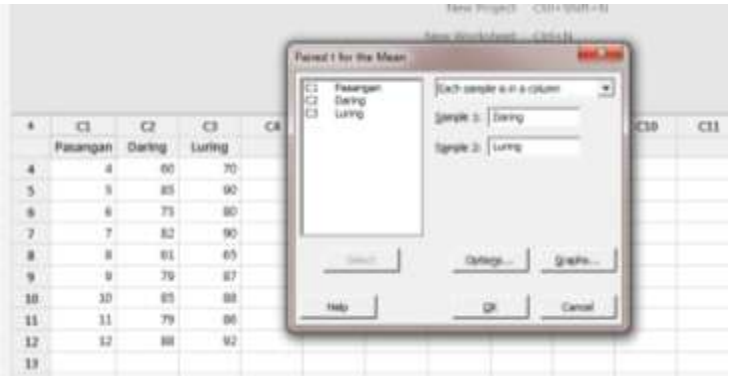
Gambar 5.2.5.2 Input data Contoh 8 pada *worksheet*

Langkah 2. Dari ikon *Stat* pilih *Basic Statistics*, kemudian *Paired* t (Gambar 5.2.5.3).



Gambar 5.2.5.3 Menu *Paired* t (pengamatan berpasangan)

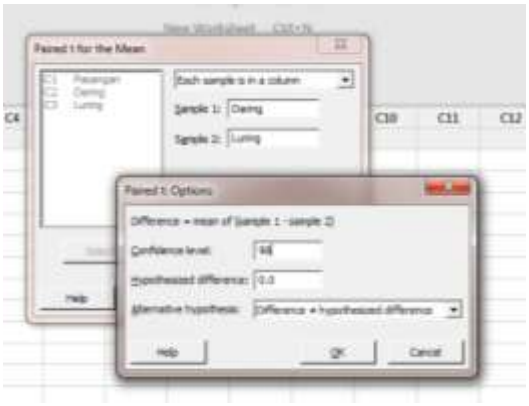
Langkah 3. Dari Langkah 2 akan muncul kotak *Paired* t *for the Mean*. Pilih *Each sample is in a column*, kemudian input data Daring ke *Sample*1 dan Data Luring ke *Sample*2 (Gambar 5.2.5.4).



Gambar 5.2.5.4 Kotak dialog *Paired* t *for the Mean*

Kemudian klik ikon *Options*.

Langkah 4. Akan muncul kotak dialog *Paired* t: *Options*. Isikan *Confidence level* 98. *Hypothesized difference* dan *Alternative hypothesis* sesuai *d*efault. Kemudian klik Ok sebanyak dua kali (Gambar 5.2.5.5).



Gambar 5.2.5.5 Kotak dialog *Paired* t: *Options*

Hasilnya diberikan pada Gambar 5.2.5.6 berikut.

Gambar 5.2.5.6 Hasil pengujian hipotesis Contoh 8

Dari Gambar 5.2.5.6 diperoleh:

a. Selang kepercayaan 98% untuk rata-rata selisih nilai ujian kedua cara belajar adalah,

$$-8{,}05 < \mu_D < -2{,}12$$

Selang kepercayaan ini tidak memuat nilai selisih rata-rata yang dihipotesiskan yaitu,

$$H_0 : \mu_D = 0$$

Dari selang kepercayaan ini kita sudah bisa menyimpulkan bahwa H_0 ditolak.

b. T-*value* = –4,66 dengan P-*value* = 0,001 < α = 0,02.

Maka H_0 ditolak bahwa rata-rata selisih nilai ujian kedua cara belajar adalah tidak sama dengan nol, dengan perkataan lain rata-rata selisih nilai ujian cara belajar daring dan luring adalah tidak sama

5.2.6. Uji hipotesis variansi suatu populasi

Contoh 9.

Data berikut menyatakan panjang, dalam cm, 10 bayi perempuan yang baru lahir di sebuah rumahsakit : 49 ; 48 ; 50 ; 49,0 ; 49 ; 51 ; 52 ; 49 ; 49 ; dan 50. Ujilah hipotesis yang menyatakan bahwa variansi panjang semua bayi yang lahir di rumahsakit tersebut adalah 0,5, dengan hipotesis alternatif tidak sama dengan 0,5. Gunakan α = 5%.

Penyelesaian

Langkah pengujian hipotesis.

1. $H_0 : \sigma^2 = 0{,}5$
2. $H_1 : \sigma^2 \neq 0{,}5$
3. $\alpha = 0{,}05$
4. Dari Tabel 5.2.1 gunakan uji statistik nomor 7 yaitu,

$$X^2 = \frac{(n-1)S^2}{\sigma_0^2} \quad , \quad \nu = n - 1 = 9.$$

Berdasarkan H_1 maka daerah kritis adalah ujung kiri dan ujung kanan yaitu $X^2 < \chi^2_{1-\alpha/2}$ dan $X^2 > \chi^2_{\alpha/2}$. α = 0,05 $\Rightarrow$ $\alpha/2$ = 0,025 dan dengan menggunakan Program Minitab (Tabel 5.2.6.1 dan Tabel 5.2.6.2).

Tabel 5.2.6.1 Nilai $\chi^2_{0,025,\nu=9}$

P(X ≤ x)	x
0.975	19.0228

Tabel 5.2.6.2 Nilai $\chi^2_{0,975,\nu=9}$

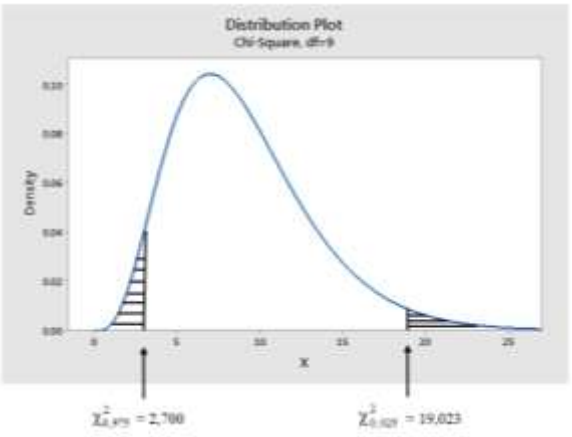
P(X ≤ x)	x
0.025	2.70039

Maka:

$$\chi^2_{1-\alpha/2} = \chi^2_{0,975} = 2{,}700 \text{ dan } \chi^2_{\alpha/2} = \chi^2_{0,025} = 19{,}023$$

Dan daerah kritis:

$$X^2 < 2{,}700 \text{ dan } X^2 > 19{,}023 \text{ (Gambar 5.2.6.1)}$$



Gambar 5.2.6.1 Daerah kritis Contoh 9

5. Dari Contoh 4.7.1 halaman 87, dengan Program Minitab diperoleh $s^2 = 1{,}38$

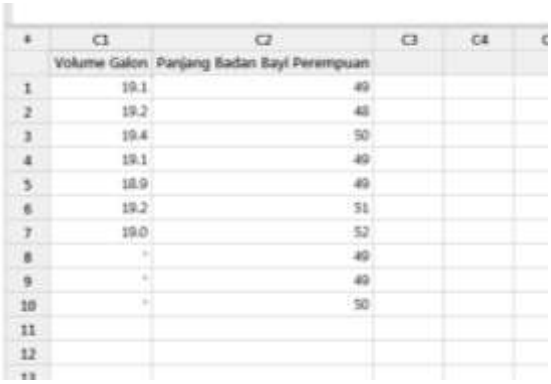
   Maka nilai statistik,

$$X^2_{sampel} = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(10-1).(1{,}38)}{0{,}5} = 24{,}8$$

6. Kesimpulan : X^2_{sampel} terletak pada daerah kritis di sebelah kanan, maka H_0 ditolak dan kita menerima H_1 yaitu $\sigma^2 \neq 0{,}5$.

Berikut langkah pengujian hipotesis dengan menggunakan Program Minitab.

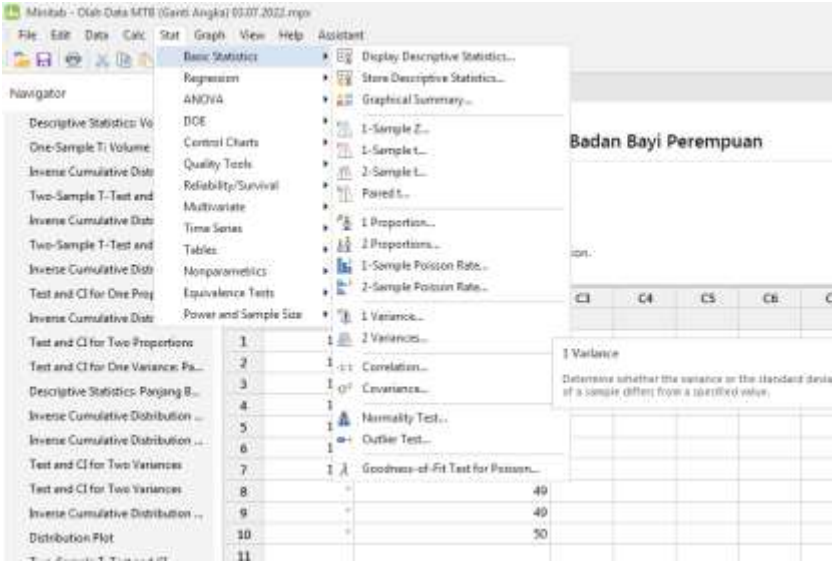
Langkah 1. Input data pada lembar kerja (*worksheet*) seperti yang diperlihatkan pada Gambar 5.2.6.2 berikut.



	C1	C2	C3	C4
	Volume Galon	Panjang Badan Bayi Perempuan		
1	19.1	49		
2	19.2	48		
3	19.4	50		
4	19.1	49		
5	18.9	49		
6	19.2	51		
7	19.0	52		
8	*	49		
9	*	49		
10	*	50		
11				
12				

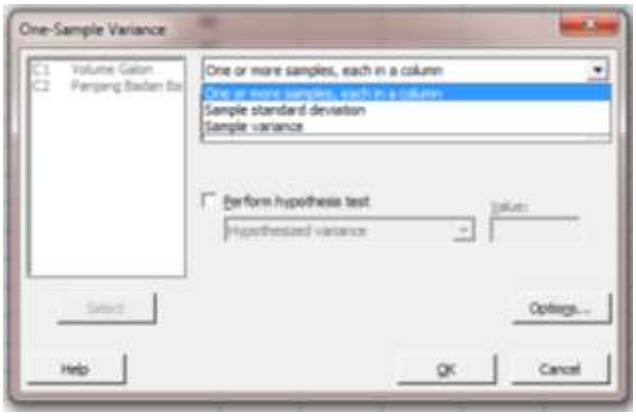
Gambar 5.2.6.2 Input data Contoh 9

Langkah 2. Dari ikon *Stat* pilih *Basic Statistics*, kemudian 1 *Variance* (Gambar 5.2.6.3).



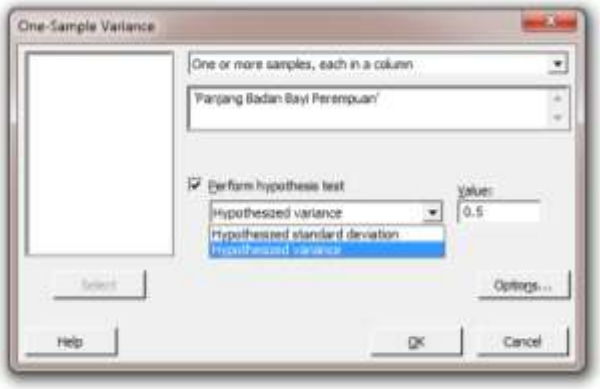
Gambar 5.2.6.3 Menu 1 *Variance*

Langkah 3. Dari Langkah 2 akan muncul kotak dialog *One-Sample Variance*. Pilih *One or more samples, each in a column* (ini *default* pilihan) (Gambar 5.2.6.4).



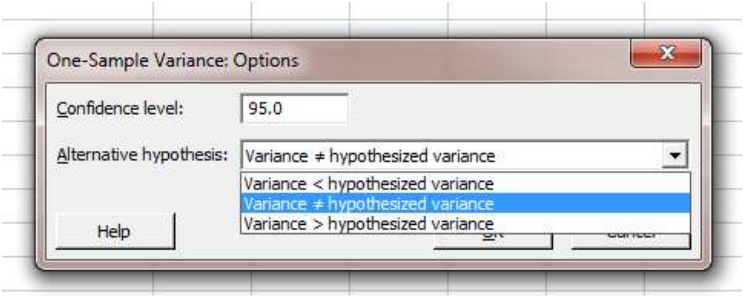
Gambar 5.2.6.4 Kotak dialog *One-Sample Variance*

Langkah 4. Kemudian input data C2 Panjang Badan Bayi Perempuan ke dalam kotak data, centang *Perform hypothesis test.* Pilih *Hypothesized variance* (ini adalah *default* pilihan). Isikan *Value 0,5* (Gambar 5.2.6.5).



Gambar 5.2.6.5 Kotak dialog *Perform hypothesis test One-Sample Variance*

Langkah 5. Klik ikon *Options*. Akan muncul kotak dialog *One-Sample Variance*: *Options*. Isikan *Confidence level* 95. Pilih *Alternative hypothesis* sesuai *default* yaitu *Variance ≠ hypothesized variance*. Kemudian klik Ok sebanyak dua kali (Gambar 5.2.6.6).



Gambar 5.2.6.6 Kotak dialog *One-Sample Variance*: *Options*

Hasilnya diberikan pada Gambar 5.2.6.7 berikut.

Test and CI for One Variance: Panjang Badan Bayi Perempuan

Method

σ: standard deviation of Panjang Badan Bayi Perempuan
The Bonett method is valid for any continuous distribution.
The chi-square method is valid only for the normal distribution.

Descriptive Statistics

N	StDev	Variance	95% CI for σ using Bonett	95% CI for σ using Chi-Square
10	1.17	1.38	(0.62, 2.75)	(0.81, 2.14)

Test

Null hypothesis $H_0: \sigma^2 = 0.5$
Alternative hypothesis $H_1: \sigma^2 \neq 0.5$

Method	Test Statistic	DF	P-Value
Bonett	—	—	0.102
Chi-Square	24.80	9	0.006

Gambar 5.2.6.7 Hasil pengujian hipotesis Contoh 9

Dari Gambar 5.2.6.7 diperoleh:

a. Selang kepercayaan 95% untuk standar deviasi Panjang Badan Bayi Perempuan (*Chi-square* atau Khi-kuadrat) adalah,

$$0{,}81 < \sigma < 2{,}14$$

Maka dengan mengkuadratkan semua ruas dalam pertekasamaan di atas diperoleh selang kepercayaan 95% untuk variansi Panjang Badan Bayi Perempuan adalah,

$$0{,}65 < \sigma^2 < 4{,}6$$

Selang kepercayaan ini tidak memuat nilai variansi yang dihipotesiskan yaitu,

$$H_0 : \sigma^2 = 0{,}5$$

Dari selang kepercayaan ini kita sudah bisa menyimpulkan bahwa H_0 ditolak.

b. $X^2_{sampel} = 24{,}8$ dengan P-*value* = 0,006 < α = 0,05.

Maka H_0 ditolak bahwa variansi Panjang Badan Bayi Perempuan tidak sama dengan 0,5.

5.2.7. Uji hipotesis rasio variansi dua populasi

Contoh 10.

Seorang Dosen Penasehat Akademik di suatu perguruan tinggi ingin membandingkan variansi nilai matakuliah Matematika I dan Fisika Dasar untuk mahasiswa semester I. Diketahui bahwa rataan dan variansi dari 20 peserta matakuliah Matematika I masing-masing adalah 80 dan 7.

Rataan dan standar deviasi dari 25 peserta matakuliah Fisika Dasar masing-masing adalah 70 dan

5. Ujilah hipotesis bahwa $\frac{\sigma_1^2}{\sigma_2^2}=1$ lawan alternatif $\frac{\sigma_1^2}{\sigma_2^2}\neq 1$; bila σ_1^2 dan σ_2^2 masing-masing menyatakan variansi populasi nilai matakuliah Matematika I dan Fisika Dasar.

Penyelesaian

Langkah pengujian hipotesis.

1. $H_0 : \frac{\sigma_1^2}{\sigma_2^2}=1$
2. $H_1 : \frac{\sigma_1^2}{\sigma_2^2}\neq 1$
3. $\alpha = 0{,}02$
4. Dari Tabel 5.2.1 gunakan uji statistik nomor 7 yaitu,

   $F = S_1^2 / S_2^2$ dengan $\nu_1 = n_1 - 1 = 20 - 1 = 19$ dan $\nu_2 = n_2 - 1 = 25 - 1 = 24$

   Berdasarkan H_1 maka daerah kritis adalah ujung kiri dan ujung kanan yaitu $F < f_{1-\alpha/2}(\nu_1, \nu_2)$ dan $F > f_{\alpha/2}(\nu_1, \nu_2)$, $\alpha = 0{,}02 \Rightarrow \alpha/2 = 0{,}01$ dan dengan menggunakan Program Minitab:

   Tabel 5.2.7.1 Nilai $f_{0{,}01,\,(19,\,24)}$

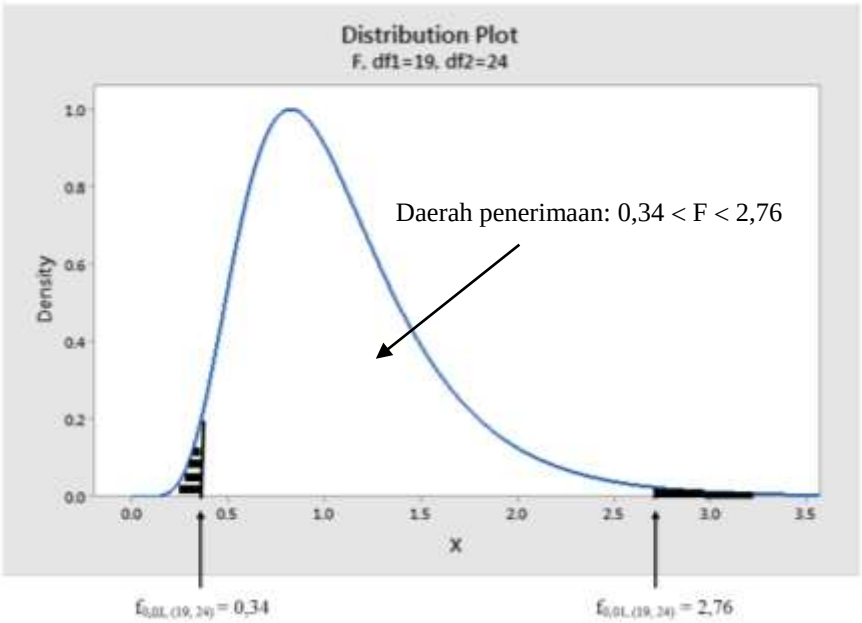
P(X ≤ x)	x
0,99	2,76239

   Tabel 5.2.7.2 Nilai $f_{0{,}99,\,(19,\,24)}$

P(X ≤ x)	x
0,99	0,341896

   Maka daerah kritis adalah $F < 0{,}34$ dan $F > 2{,}76$ (Gambar 5.2.7.1).
5. Nilai statistik uji:

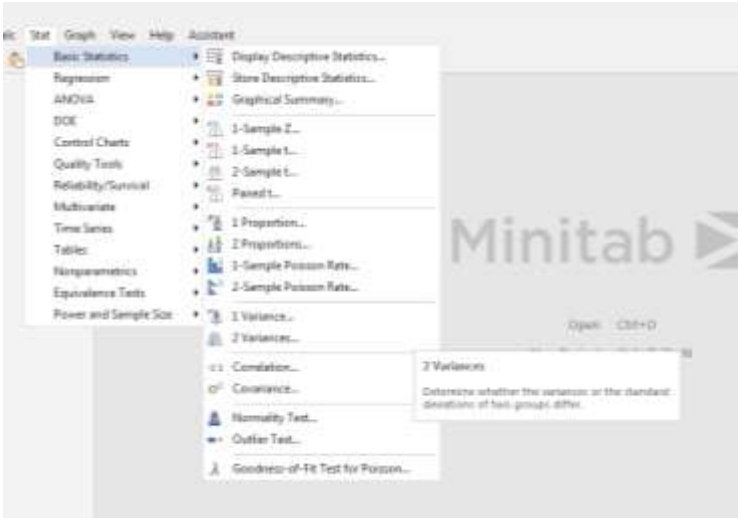
   $f_{sampel} = s_1^2 / s_2^2 = 49/25 = 1{,}96$
6. Kesimpulan : Nilai f_{sampel} terletak pada daerah penerimaan yaitu selang $0{,}34 < F < 2{,}76$. Maka hipotesis nol diterima bahwa $\frac{\sigma_1^2}{\sigma_2^2}=1$ atau $\sigma_1^2=\sigma_2^2$ yaitu variansi populasi nilai matakuliah Matematika I sama dengan variansi nilai Fisika Dasar.



Gambar 5.2.7.1 Daerah kritis Contoh 10

Berikut langkah pengujian hipotesis dengan menggunakan Program Minitab.

Langkah 1. Karena data pada contoh ini merupakan ringkasan data, maka kita langsung ke ikon *Stat* pilih *Basic Statistics*, kemudian 2 *Variance* (Gambar 5.2.7.2).

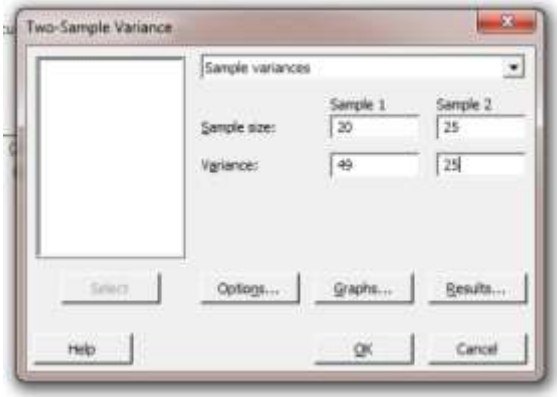


Gambar 5.2.7.2 Menu 2 *Variance*

Langkah 2. Akan muncul kotak dialog *Two-Sample Variance*. Karena kita akan menguji rasio variansi, maka pilih *Sample variance* (Gambar 5.2.7.3).

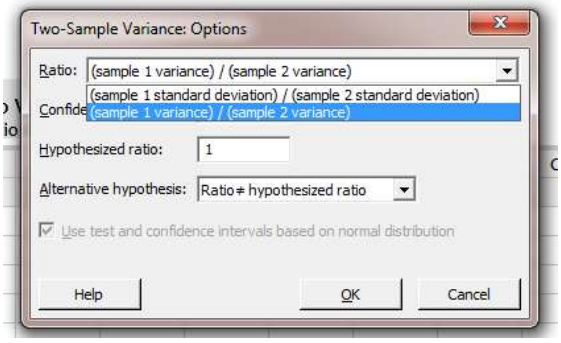
Gambar 5.2.7.3 Kotak dialog *Two-Sample Variance*

Kemudian isikan data sesuai data yang diberikan pada Contoh 10 seperti yang diberikan pada Gambar 5.2.7.4.



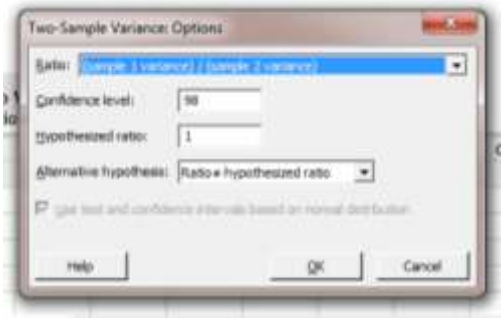
Gambar 5.2.7.4 Input data Contoh 10

Langkah 3. Klik ikon *Options*. Akan muncul kotak dialog *Two-Sample Variances* : *Options*. Untuk *Ratio* pilih (*sample 1 variance* / *sample 2 variance*) (Gambar 5.2.7.5).



Gambar 5.2.7.5 Kotak dialog *Two-Sample Variances* : *Options* (1)

Kemudian isikan *Confidence level* 98, dan *Hypothesized ratio* = 1 (ini *default Ratio*) (Gambar 5.2.7.6)



Gambar 5.2.7.6 Kotak dialog *Two-Sample Variances* : *Options* (2)

Langkah 5. Klik Ok dua kali dan hasilnya diberikan pada Gambar 5.2.7.7.

Ratio of Variances

Estimated Ratio	98% CI for Ratio using F
1.96	(0.710, 5.733)

Test

Null hypothesis	$H_0: \sigma_1^2 / \sigma_2^2 = 1$
Alternative hypothesis	$H_1: \sigma_1^2 / \sigma_2^2 \neq 1$
Significance level	$\alpha = 0.02$

Method	Test Statistic	DF1	DF2	P-Value
F	1.96	19	24	0.120

Gambar 5.2.7.7 Hasil pengujian hipotesis Contoh 10

Dari Gambar 5.2.7.7 diperoleh:

a. Selang kepercayaan 98% untuk rasio variansi $\frac{\sigma_1^2}{\sigma_2^2}$ adalah,

$$0{,}710 < \frac{\sigma_1^2}{\sigma_2^2} < 5{,}733$$

Selang kepercayaan ini memuat nilai variansi yang dihipotesiskan yaitu,

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$$

Dari selang kepercayaan ini kita sudah bisa menyimpulkan bahwa H_0 diterima.

b. Nilai $f_{sampel} = 1{,}96$ dengan P-*value* $= 0{,}12 > \alpha = 0{,}02$.

Maka H_0 diterima bahwa $\frac{\sigma_1^2}{\sigma_2^2} = 1$ atau $\sigma_1^2 = \sigma_2^2$ yaitu variansi populasi nilai matakuliah Matematika I sama dengan variansi nilai Fisika Dasar.

5.3. Uji Hipotesis Proporsi

5.3.1. Uji hipotesis proporsi suatu populasi

Uji menyangkut proporsi merupakan uji yang berkaitan dengan peubah acak Binomial yang merupakan peubah acak diskrit. Dalam masalah pengujian hipotesis yang menyangkut proporsi ini, prosedur pengujian hipotesis tergantung pada ukuran sampel dalam percobaan, yaitu:

a. Apabila ukuran sampel kecil, digunakan Distribusi Binomial dengan langkah pengujian hipotesis diberikan sebagai berikut.

   Langkah pengujian hipotesis.

   1. $H_0 : p = p_0$
   2. $H_1 : p < p_0$ atau $p > p_0$ atau $p \neq p_0$
   3. Pilih taraf keberartian α
   4. Uji statistik menggunakan Distribusi Binomial
   5. Daerah Kritis:
      a. $H_1 : p < p_0$

         Semua nilai x sehingga $P(X \leq x \mid p = p_0) < \alpha$

         Nilai-p = $P(X \leq x \mid p = p_0)$

      b. $H_1 : p > p_0$

         Semua nilai x sehingga $P(X \geq x \mid p = p_0) > \alpha$

         Nilai-p = $P(X \geq x \mid p = p_0)$

      c. $H_1 : p \neq p_0$

         (i) Semua nilai x sehingga $P(X \leq x \mid p = p_0) < \alpha/2$ bila $x < np_0$

         Nilai-p = $2\ P(X \leq x \mid p = p_0)$

         (ii) Semua nilai x sehingga $P(X \geq x \mid p = p_0) < \alpha/2$ bila $x > np_0$

         Nilai-p = $2\ P(X \geq x \mid p = p_0)$

      Catatan. Karena X adalah peubah acak diskrit, maka jarang dapat dibuat daerah kritis yang tepat, oleh karenanya dalam perhitungan nilai-p kadang-kadang mungkin akan terjadi sedikit perbedaan antara nilai yang dihitung secara manual dengan yang diolah menggunakan Program Minitab.

   6. Perhitungan : tentukan x dan hitung peluang yang sesuai
   7. Kesimpulan : tolak H_0 bila x jatuh dalam daerah kritis ; kalau tidak, terima H_0.

Contoh 11

Pada kasus pasien penderita penyakit tertentu yang sulit sembuh di suatu rumah sakit. Diketahui bahwa 4 dari 15 pasien yang sembuh; di mana ke-15 pasien tersebut tidak mempunyai hubungan darah (saling bebas). Ujilah hipotesis bahwa proprosi sesungguhnya dari seluruh pasien yang menderita penyakit tersebut adalah 0,6, dengan hipotesis alternatif lebih kecil dari 0,6. Gunakan taraf signifikansi 0,05.

Penyelesaian

Dalam masalah ini, proporsi populasi yaitu $p_0 = 0,6$ dan $x = 4$ dan $n = 15$.

1. H_0 : $p = 0,6$
2. H_1 : $p < 0,6$
3. $\alpha = 0,05$
4. Uji statistik menggunakan Distribusi Binomial
5. Daerah Kritis:

   H_1 : $p < 0,6$

   Semua nilai x sehingga $P(X \leq x \mid p = 0,6) < \alpha$

   Maka kita hitung dengan Program Minitab: $P(X \leq 4 \mid p = 0,6) = 0,009$ (Tabel)

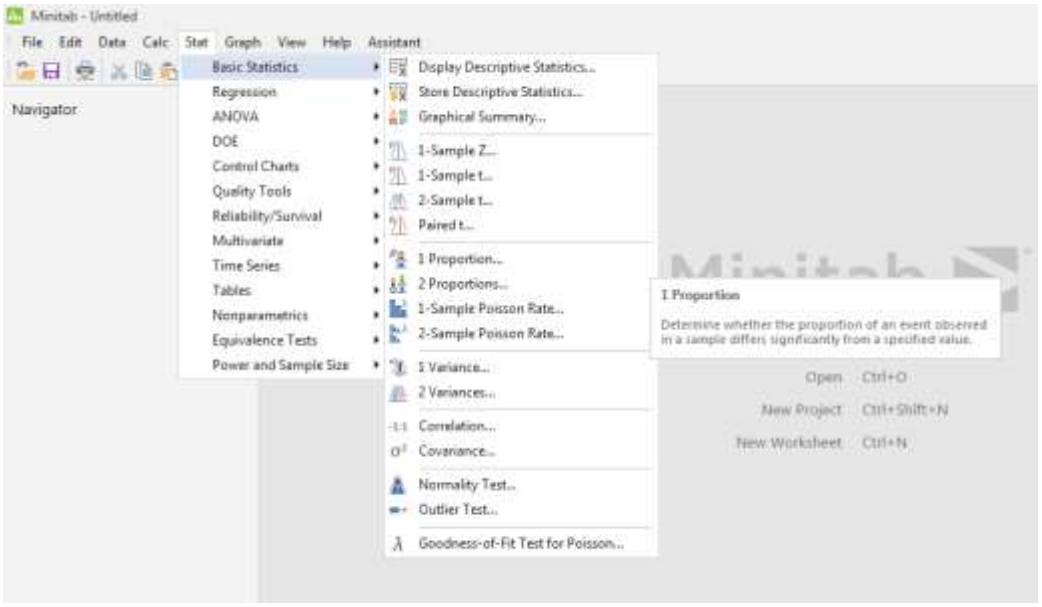
Tabel 5.3.1.1 Nilai peluang Binomial $P(X \leq 4 \mid p = 0,6, n = 15)$

x	P(X ≤ x)
4	0,0093477

6. Kesimpulan : tolak H_0 karena $P(X \leq 4 \mid p = 0,6) = 0,009 < 0,05$.

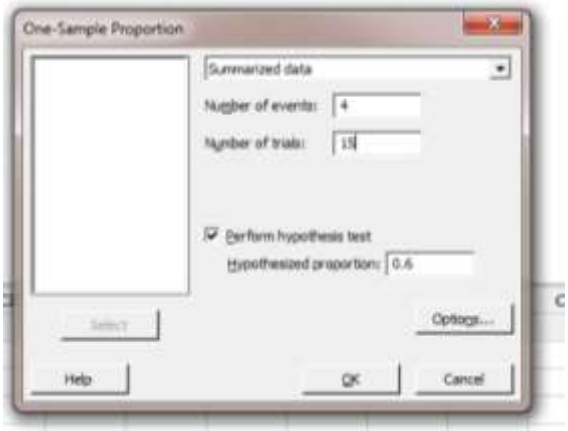
Bwrikut ini langkah pengujian hipotesis dengan menggunakan Program Minitab.

Langkah 1. Dari Menu Stat, pilih Basic Statistics, kemudian klik 1 *Proportion* (Gambar 5.3.1.1).



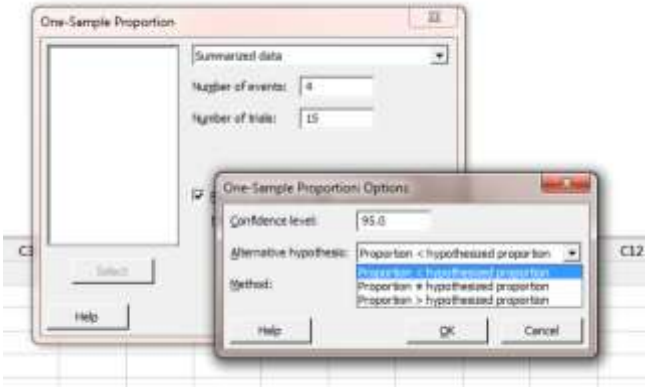
Gambar 5.3.1.1 Menu 1 *Proportion*

Langkah 2. Setelah muncul kotak dialog *One-Sample Proportion* pilih *Summarized data* kemudian isikan data-data sesuai data pada Contoh 11 yaitu *Number of events* 4 dan *Number of trials* 15. Lalu centang *Perform hypothesis test* dan isikan data *Hypothesized proportion* 0,6. Lalu klik *Options* (Gambar 5.3.1.2).



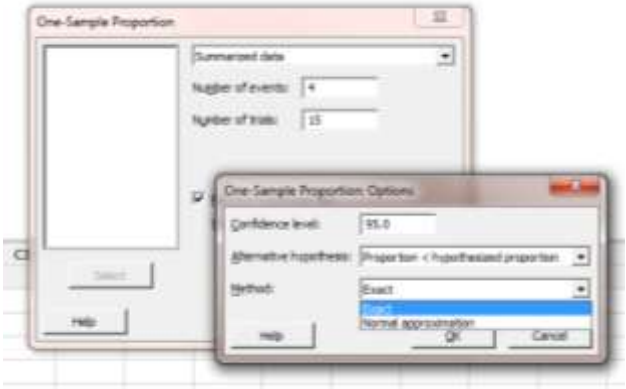
Gambar 5.3.1.2 Kotak dialog *One-Sample Proportion*

Langkah 3. Akan muncul kotak dialog *One-Sample Proportion* : *Options*. Data *Confidence level* sudah sesuai *default* yaitu 95. Pada *Alternative hypothesized* pilih *Proportion* < *hypothesized proportion* (Gambar 5.3.1.3).



Gambar 5.3.1.3 Kotak dialog *One-Sample Proportion* : *Options* (1)

Langkah 4. Pada *Method* pilih *Exact*. Ini merupakan pilihan apabila ukuran sampel kecil (Gambar 5.3.1.4).



Gambar 5.3.1.4 Kotak dialog *One-Sample Proportion* : *Options* (2)

Langkah 5. Klik Ok dua kali dan hasilnya diberikan pada Gambar 5.3.1.5 berikut.

Test and CI for One Proportion

Method

p: event proportion
Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% Upper Bound for p
15	4	0.266667	0.510752

Test

Null hypothesis	H_0: p = 0.6
Alternative hypothesis	H_1: p < 0.6

P-Value
0.009

Gambar 5.3.1.5 Hasil pengujian hipotesis Contoh 11 (1)

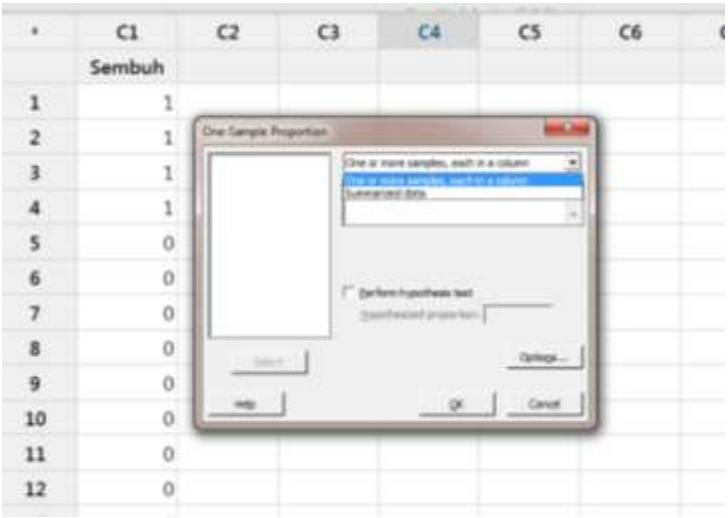
Dari Gambar 5.3.1.5 di atas terlihat diperoleh nilai-p = 0,009, di mana hasil ini sama dengan nilai-p yang dihitung secara manual. Dengan demikian diperoleh kesimpulan yang sama yaitu hipotesis nol ditolak dan kita menerima H_1 bahwa proporsi pasien yang sembuh dari seluruh pasien yang mungkin menderita penyakit tersebut lebih kecil dari 0,6.

Berikut ini cara lain untuk melakukan prosedur pengujian hipotesis pada contoh di atas. Perbedaannya adalah cara menginput data. Dalam contoh ini, kemungkinan ada 2 yaitu pasien sembuh (diberi kode "1") dan pasien tidak sembuh (kode "0"). Input data pada kolom C1 dan beri nama Sembuh seperti pada Gambar 5.3.1.6 berikut.

	C1	C2	C3	C4	C5	C6
	Sembuh					
1	1					
2	1					
3	1					
4	1					
5	0					
6	0					
7	0					
8	0					
9	0					
10	0					
11	0					
12	0					
13	0					
14	0					
15	0					

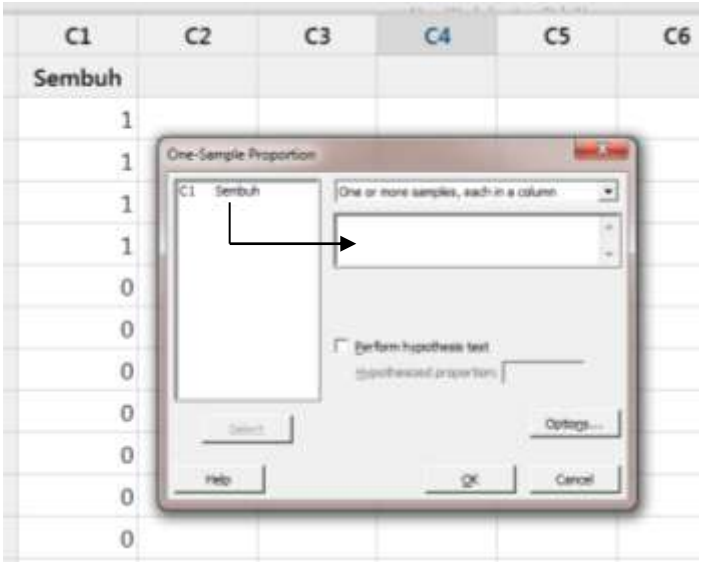
Gambar 5.3.1.6 Input data Contoh 1 pada *worksheet*

Pada kotak dialog *One-Sample Proportion* pilih *One or more samples, each in a column* (Gambar 5.3.1.7).



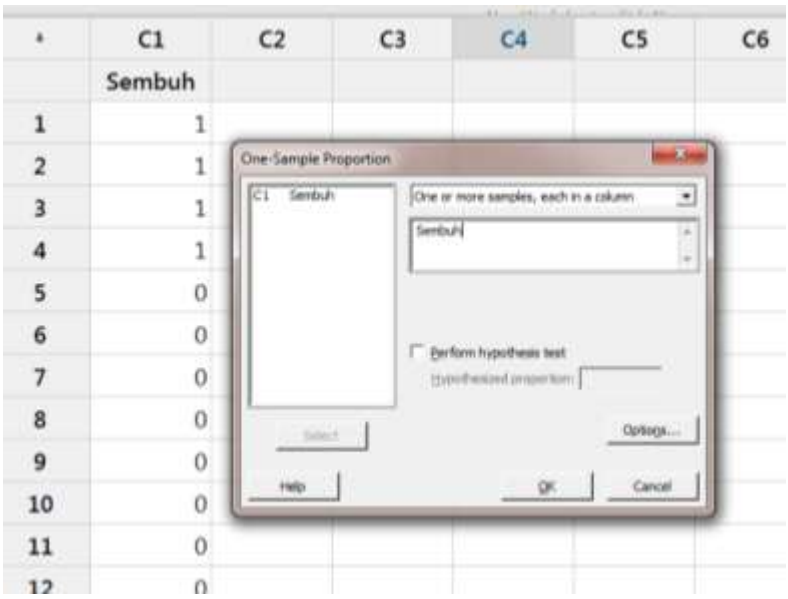
Gambar 5.3.1.7 Opsi *One or more samples, each in a column*

Kemudian masukkan data C1 Sembuh ke kotak data (Gambar 5.3.1.8).



Gambar 5.3.1.8 Input data dari kotak data

Setelah data pada kolom C1 Sembuh masuk ke kotak data (Gambar 5.3.1.9), langkah selanjutnya sama seperti sebelumnya (Gambar 5.3.1.2 dan seterusnya).



Gambar 5.3.1.9 Data pada kolom C1 Sembuh masuk ke kotak data

Hasilnya diberikan pada Gambar 5.3.1.10 berikut.

Test and CI for One Proportion: Sembuh

Method

Event: Sembuh = 1
p: proportion where Sembuh = 1
Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% Upper Bound for p
15	4	0.266667	0.510752

Test

Null hypothesis	H_0: p = 0.6
Alternative hypothesis	H_1: p < 0.6

P-Value
0.009

Gambar 5.3.1.10 Hasil pengujian hipotesis Contoh 11 (2)

Pada bagian atas Gambar di atas terlihat ad penjelasan tentang *Event*: Sembuh = 1 yang menjadi peubah acak Binomial pada contoh ini.

Berikut ini diberikan hasil pengujian hipotesis jika dipilih hipotesis alternatif $H_1 : p \neq 0{,}6$ (Gambar 5.3.1.11).

Test and CI for One Proportion: Sembuh

Method

Event: Sembuh = 1
p: proportion where Sembuh = 1
Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% CI for p
15	4	0.266667	(0.077872, 0.551003)

Test

Null hypothesis	H_0: p = 0.6
Alternative hypothesis	H_1: p ≠ 0.6

P-Value
0.015

Gambar 5.3.1.11 Hasil pengujian hipotesis Contoh 11 (3)

Dalam contoh ini $np_0 = (15)\,(0{,}6) = 9 \rightarrow x = 4 < np_0 = 9$.

Maka:

$$\text{Nilai-p} = 2\,P(X \leq 4 \mid p = 0{,}6) = (2)\,(0{,}009) = 0{,}018$$

Sedangkan dari Gambar 5.3.1.11 di atas diperoleh nilai-p = 0,015. Di sini terdapat sedikit perbedaan yang tidak terlalu besar yaitu 0,003. Ini merupakan salah satu hal yang mungkin terjadi karena X merupakan peubah acak diskrit, sehingga terkadang sulit menentukan nilai-p yang tepat.

Contoh 12

Pada kasus pasien penderita penyakit tertentu yang sulit sembuh di suatu rumah sakit, diketahui bahwa 4 dari 15 pasien yang sembuh; di mana ke-15 pasien tersebut tidak mempunyai hubungan darah (saling bebas). Ujilah hipotesis bahwa proprosi sesungguhnya dari seluruh pasien yang menderita penyakit tersebut adalah 0,1 dengan hipotesis alternatif lebih besar dari 0,1. Gunakan taraf signifikansi 0,05.

Penyelesaian

Dalam masalah ini, proporsi populasi yaitu $p_0 = 0,05$ dan x = 4 dan n = 15.

1. H_0 : p = 0,1
2. H_1 : p > 0,1
3. $\alpha = 0,05$
4. Uji statistik menggunakan Distribusi Binomial
5. Daerah Kritis:

   H_1 : p > 0,1

   Semua nilai x sehingga $P(X \geq x \mid p = 0,1) < \alpha$

   Maka kita hitung dengan Program Minitab:

   $P(X \geq 4 \mid p = 0,1) = 1 - P(X \leq 3 \mid p = 0,1) = 1 - 0,94 = 0,06$ (Tabel)

Tabel 5.3.1.2 Nilai peluang Binomial: $P(X \leq 3 \mid p = 0,1, n = 15)$

x	$P(X \leq x)$
3	0,94

6. Kesimpulan : tolak H_0 karena $P(X \geq 4 \mid p = 0,1) = 0,06 > 0,05$.

Pada Gambar 5.3.1.12 berikut diberikan hasil pengolahan data apabila H_0 : p = 0,1 dan H_1 : p > 0,1 (diberikan tanpa penjelasan penggunaan Program Minitab).

Test and CI for One Proportion: Sembuh

Method

Event: Sembuh = 1
p: proportion where Sembuh = 1
Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% Lower Bound for p
15	4	0.266667	0.096658

Test

Null hypothesis	H_0: p = 0.1
Alternative hypothesis	H_1: p > 0.1

P-Value
0.056

Gambar 5.3.1.12 Hasil pengujian hipotesis Contoh 12

b. Apabila ukuran sampel besar, digunakan pendekatan Distribusi Normal (Z) terhadap Distribusi Binomial dengan langkah pengujian hipotesis yang sama seperti sebelumnya. Kita akan membahas Contoh pada halaman dengan penyelesaian menggunakan Program Minitab.

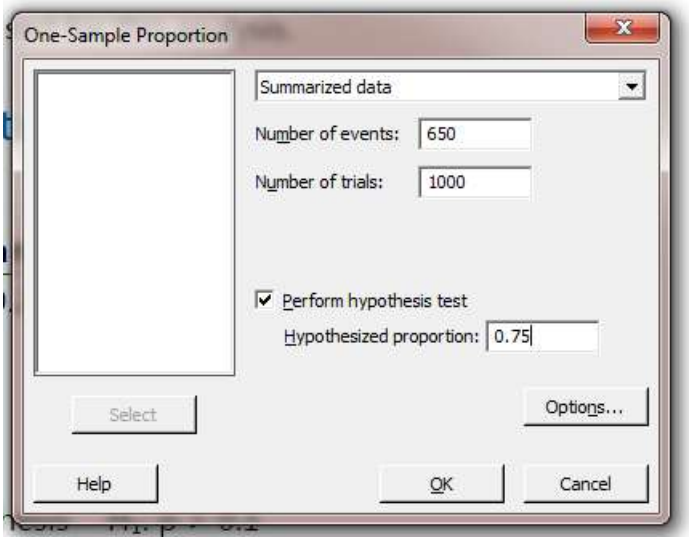
Contoh 13 (Contoh 4.6.2 halaman 85)

Diketahui bahwa dari sampel acak 1000 keluarga di suatu kota besar pada suatu provinsi ditemukan bahwa 650 keluarga menggunakan jaringan nirkabel WiFi di rumahnya. Ujilah hipotesis yang menyatakan bahwa proporsi keluarga yang menggunakan jaringan nirkabel WiFi di rumahnya di seluruh provinsi tersebut adalah 0,75. Pilih hipotesis alternatif bahwa proporsi keluarga yang menggunakan jaringan nirkabel WiFi di rumahnya di seluruh provinsi tersebut tidak sama dengan 0,75. Gunakan taraf keberartian $\alpha = 0{,}05$.

Penyelesaian

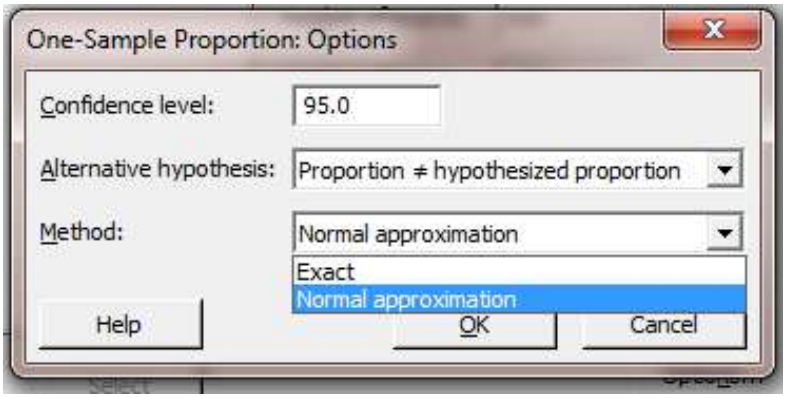
Dalam Contoh ini karena ukuran sampel besar maka akan digunakan pendekatan Distribusi Normal terhadap Distribusi Binomial. Berikut langkah pengujian dengan menggunakan Program Minitab yang caranya sama dengan yang sudah dijelaskan sebelumnya.

Langkah 1. Pada kotak dialog *One-Sample Proportion*, pilih *Summarized Data*. Kemudian isikan data-data sesuai data pada contoh yaitu *Number of events*: 650 dan *Number of trials*: 1000. Centang *Perform hypothesis test* dan isikan data *Hypothesized proportion*: 0,75. Selanjutnya klik *Options* (Gambar 5.3.1.13).



Gambar 5.3.1.13 Input data Contoh 13

Langkah 2. Setelah muncul kotak dialog *One-Sample Proportion*: *Options*, isikan lagi data-data sesuai data pada contoh. *Confidence level* sudah sesuai *default* yaitu 95. Pada *Alternative hypothesis* pilih *Proportion ≠ hypothesized proportion* dan *Method*: *Normal approximation* (Gambar 5.3.1.14).



Gambar 5.3.1.14 Kotak dialog *One-Sample Proportion*: *Options*

Langkah 3. Klik Ok dua kali dan hasilnya diberikan dalam Gambar 5.3.1.15 berikut.

Test and CI for One Proportion

Method

p: event proportion
Normal approximation method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% CI for p
1000	650	0.650000	(0.620438, 0.679562)

Test

Null hypothesis	H_0: p = 0.75
Alternative hypothesis	H_1: p ≠ 0.75

Z-Value	P-Value
-7.30	0.000

Gambar 5.3.1.15 Hasil pengolahan data Contoh 13

Dari Gambar di atas diperoleh bahwa nilai-Z = –7,3 dan nilai-p = 0,000 < α = 0,05. Maka kesimpulannya hipotesis nol ditolak dan hipotesis alternatif diterima bahwa proporsi keluarga yang menggunakan jaringan nirkabel WiFi di rumahnya di seluruh provinsi tersebut tidak sama dengan 0,75.

5.3.2. Uji hipotesis selisih proporsi dua populasi

Konsep pengujian selisih proporsi dua populasi berkaitan dengan pembahasan topik selang kepercayaan $(1 - \alpha)100\%$ untuk selisih proporsi, $p_1 - p_2$ pada halaman 84. Dari dua populasi bebas yang masing-masing berasal dari peubah Distribusi Binomial yaitu,

$$X \sim b(x; n_1, p_1) \text{ dan } Y \sim b(y; n_2, p_2)$$

maka menurut konsep TLP halaman 48 maka bentuk limit distribusi,

$$Z_1 = \frac{X - n_1 p_1}{\sqrt{n_1 p_1 (1 - p_1)}} = \frac{p_1 - p_1}{\sqrt{p_1 (1 - p_1)/n_1}} \text{ dan}$$

$$Z_2 = \frac{Y - n_2 p_2}{\sqrt{n_2 p_2 (1 - p_2)}} = \frac{p_2 - p_2}{\sqrt{\pi_2 (1 - p_2)/n_2}} \qquad (5.3.2.1)$$

masing-masing berdistribusi Normal Baku n(z; 0, 1), dengan $\hat{p}_1 = X/n_1$ dan $\hat{p}_2 = Y/n_2$ masing-masing penaksir titik bagi p_1 dan p_2; X dan Y masing-masing adalah banyaknya sukses terjadi pada kedua populasi. Dari halaman juga diperoleh bahwa, $Z_1 - Z_2$ juga mempunyai Distribusi Normal dengan rataan,

$$\mu_{p_1 - p_2} = p_1 - p_2$$

dan karena kedua populasi saling bebas maka variansi,

$$\sigma^2_{p_1 - p_2} = \frac{p_1 (1 - p_1)}{n_1} + \frac{p_2 (1 - p_2)}{n_2}$$

Sehingga,

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 (1 - p_1)}{n_1} + \frac{p_2 (1 - p_2)}{n_2}}} \qquad (5.3.2.2)$$

Maka kita dapat melakukan prosedur pengujian hipotesis selisih dua proporsi dengan menggunakan distribusi Z.

Langkah pengujian hipotesis.

1. H_0 : $p_1 = p_2$
2. H_1 : $p_1 < p_2$ atau $p_1 > p_2$ atau $p_1 \neq p_2$
3. Pilih taraf keberartian α
4. Uji statistik menggunakan Distribusi Binomial
5. Daerah Kritis:
   a. H_1 : $p < p_0$

      Daerah kritis: $Z < -z_\alpha$

   b. H_1 : $p > p_0$

      Daerah kritis: $Z > z_\alpha$

   c. H_1 : $p \neq p_0$

      $Z < -z_{\alpha/2}$ dan $Z > z_{\alpha/2}$
6. Perhitungan : tentukan z_{sampel} dan hitung nilai-p.
7. Kesimpulan : tolak H_0 bila x jatuh dalam daerah kritis ; kalau tidak, terima H_0.

Contoh 14

Diketahui bahwa dalam suatu proses produksi dengan menggunakan mesin A dan mesin B akan terdapat produk yang cacat. Sampel diambil dari produksi mesin A dan mesin B untuk melihat perbedaan proporsi yang cacat dalam produksi. Diketahui bahwa 100 dari 2000 produk yang diproduksi dengan mesin A ternyata cacat, dan 90 dari 3000 produk yang diproduksi dengan mesin B ternyata cacat. Ujilah hipotesis bahwa selisih proporsi $p_1 - p_2 = 0$ atau $p_1 = p_2$; lawan alternatif $p_1 - p_2 \neq 0$ atau $p_1 \neq p_2$; dimana p_1 dan p_2 masing-masing adalah proporsi sesungguhnya produk yang cacat dari mesin A dan mesin B. Gunakan taraf keberartian α = 0,10.

Penyelesaian

Dalam soal ini:

$$\hat{p}_1 = \frac{x}{n_1} = \frac{100}{2000} = 0{,}05 \text{ dan } \hat{p}_2 = \frac{y}{n_2} = \frac{90}{3000} = 0{,}03$$

Dan taksiran titik untuk $p_1 - p_2$ adalah : 0,05 – 0,03 = 0,02.

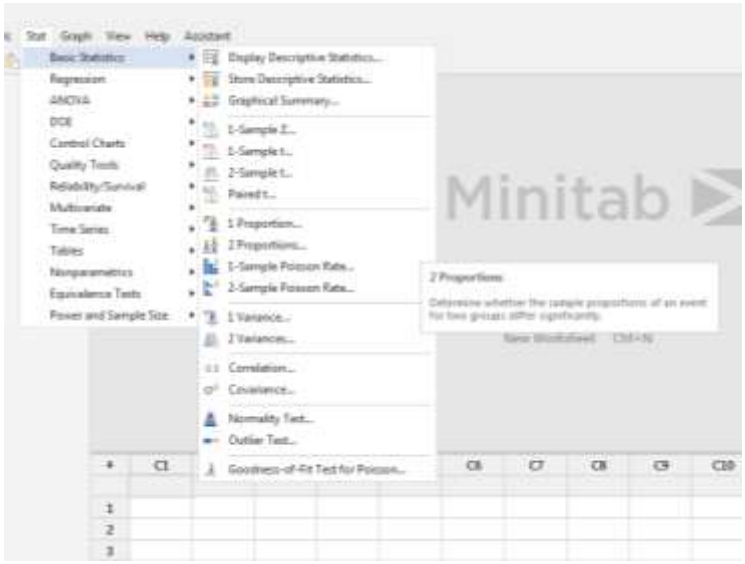
1. H_0 : $p_1 - p_2 = 0$ atau $p_1 = p_2$

2. H_1 : $p_1 - p_2 \neq 0$ atau $p_1 \neq p_2$
3. $\alpha = 0{,}10$
4. Uji statistik menggunakan Distribusi Normal
5. Daerah Kritis: $Z < -z_{\alpha/2}$ dan $Z > z_{\alpha/2}$, $\alpha = 0{,}1 \Rightarrow \alpha/2 = 0{,}05$ dan dari Contoh 4.6.2 halaman 85 diperoleh $z_{0{,}05} = 1{,}645$. Maka daerah kritis: $Z < -1{,}645$ dan $Z > 1{,}645$.
6. Perhitungan z_{sampel}.

$$Z_{sampel} = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_{11}(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{0{,}02 - 0}{\sqrt{\frac{(0{,}05)(0{,}95)}{2000} + \frac{(0{,}03)(0{,}97)}{3000}}} = 3{,}46$$

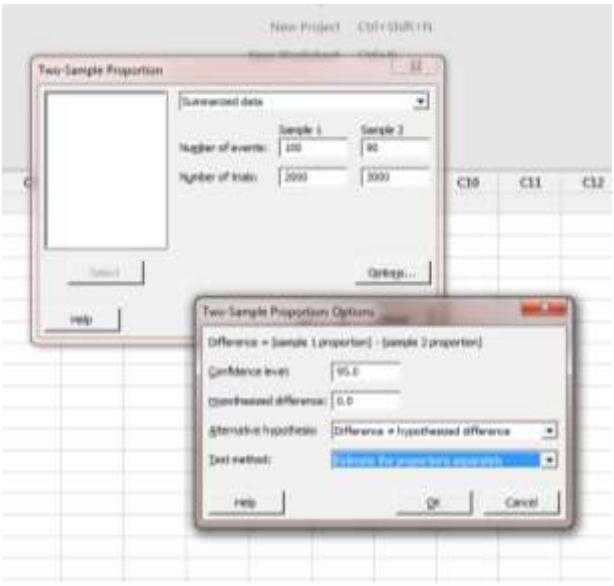
7. Kesimpulan: Tolak H_0 karena z_{sampel} terletak pada daerah kritis sebelah kanan dan terima hipotesis alternatif bahwa proporsi sesungguhnya produk yang cacat dari mesin A tidak sama dengan proporsi sesungguhnya produk yang cacat dari mesin B.

Berikut ini diberikan *output* Program Minitab untuk *Test and CI Two Proportions* Contoh 14. Gambar 5.3.2.1 memperlihatkan menu 2 *Two Proportions*.



Gambar 5.3.2.1 Menu *Two Proportions*

Pada Gambar 5.3.2.2 diberikan kotak dialog *Two-Sample Proportion*.



Gambar 5.3.2.2 Kotak dialog *Two-Sample Proportions*

Test and CI for Two Proportions

Method

p_1: proportion where Sample 1 = Event
p_2: proportion where Sample 2 = Event
Difference: $p_1 - p_2$

Descriptive Statistics

Sample	N	Event	Sample p
Sample 1	2000	100	0.050000
Sample 2	3000	90	0.030000

Estimation for Difference

Difference	90% CI for Difference
0.02	(0.010487, 0.029513)

CI based on normal approximation

Test

Null hypothesis	H_0: $p_1 - p_2 = 0$
Alternative hypothesis	H_1: $p_1 - p_2 \neq 0$

Method	Z-Value	P-Value
Normal approximation	3.46	0.001
Fisher's exact		0.000

Gambar 5.3.2.3 Hasil pengujian hipotesis Contoh 14

Hasilnya sama seperti yang dikerjakan secara manual di atas. Terlihat nilai-p dengan aproksimasi atau pendekatan Normal untuk sampel besar adalah 0,001. Maka hipotesis nol H_0 ditolak yang artinya proporsi sesunggguhnya produk yang cacat dari mesin A tidak sama dengan proporsi sesunggguhnya produk yang cacat dari mesin B .

6. Aplikasi dalam Penelitian Penulis

6.1. Sekilas Analisis Regresi

Analisis dengan Statistika Inferensi sangat banyak digunakan dalam pengolahan data suatu penelitian. Salah satu yang terkenal adalah Analisis Regresi.

Di dalam sebuah sistem dengan komponen kuantitas yang berubah-ubah, sangat menarik untuk melihat pengaruh yang ditimbulkan oleh sebagian peubah atau variabel terhadap peubah lainnya. Mungkin saja memang ada suatu hubungan fungsional yang sederhana antara peubah-peubah tersebut. Dalam hal demikian, kita mungkin ingin menghampiri atau mengaproksimasi suatu hubungan fungsional ini dengan suatu fungsi matematis yang sederhana, misalnya suatu polinom, yang mengandung peubah-peubah yang layak dan menghampiri fungsi yang sebenarnya dalam kisaran terbatas peubah-peubah tadi. Dengan mempelajari fungsi semacam ini, kita mungkin dapat memperoleh informasi lebih banyak tentang hubungan sebenarnya yang mendasarinya dan dapat memperhitungkan pengaruh terpisah dan bersama-sama yang dihasilkan oleh perubahan-perubahan pada peubah-peubah penting tertentu. Bahkan dalam hal tidak ada relasi fisis yang terpikirkan antara peubah-peubah, kita mungkin masih ingin menghubungkan peubah-peubah itu melalui suatu persamaan matematis tertentu. Meskipun persamaan itu mungkin tidak ada artinya secara fisik, namun bisa jadi persamaan semacam itu sangat bermanfaat untuk meramalkan nilai-nilai sebagian peubah tadi bila nilai-nilai peubah lainnya diketahui, mungkin di bawah kondisi tertentu yang dinyatakan. Pada tahap ini akan kita bedakan 2 (dua) jenis peubah, yaitu peubah prediktor (*predictor variable*) atau peubah bebas (*independent variable*), dan peubah respons (*response variable*) atau peubah tidak bebas (*dependent variable*).

Peubah prediktor adalah peubah yang nilainya dapat ditentukan atau diatur; biasanya di laboratorium (misalnya suhu input atau laju pemberian katalisator) atau yang nilainya dapat diamati namun tidak dapat dikendalikan, biasanya di lokasi/lapangan (misalnya kelembaban udara luar). Akibat perubahan yang disengaja, atau yang terjadi pada peubah peramal, suatu pengaruh atau efek dipancarkan terhadap peubah lain, yaitu peubah respons (misalnya warna akhir atau kemurnian suatu produk kimia tertentu).

Ada 2 jenis data dalam Analisis Regresi. Jika nilai-nilai peubah prediktor (x) dalam regresi tidak terkontrol (yaitu, tidak ditetapkan terlebih dahulu sebelum nilai peubah respon y diamati) tetapi diukur tanpa kesalahan, datanya disebut data observasi. Jika nilai-nilai peubah

bebas (x) dalam regresi dikendalikan menggunakan rancangan percobaan (yaitu, ditetapkan terlebih dahulu sebelum nilai peubah respon y diamati), datanya disebut data eksperimen.

Pada umumnya kita ingin mengetahui bagaimana perubahan pada peubah (-peubah) prediktor mempengaruhi nilai peubah respons. Pembedaan antara peubah prediktor dan peubah respons tidak selalu tampak jelas dan adakalanya tergantung pada tujuan kita. Mungkin saja pada suatu tahap sebuah peubah dianggap sebagai peubah respons, namun pada tahap lain dianggap sebagai peubah prediktor.

Jadi apabila kita menaruh perhatian pada ketergantungan suatu peubah acak Y terhadap suatu besaran atau kuantitas X yang bervariasi namun bukan merupakan peubah acak, maka suatu persamaan yang menghubungkan Y dan X disebut persamaan atau model regresi. Dalam persamaan ini, kita mengasumsikan bahwa peubah prediktor tidak bersifat acak, sedangkan peubah responsnya bersifat acak.

Bentuk umum model regresi (disebut juga model peluang/probabilitas) diberikan dalam Persamaan 6.1.1 berikut.

$$y = E(y) + \varepsilon \qquad (6.1.1)$$

di mana:

y = Variabel respon

E(y) = Komponen deterministik (*deterministic component*)

Atau disebut juga rata-rata (*mean or expected*) dari Y.

ε = Komponen galat acak (*random error component*)

Atau disebut juga galat/residu

Model regresi yang paling sederhana adalah model dengan satu perubah peramal dan grafiknya merupakan suatu garis lurus dengan kemiringan tertentu, sehingga model ini sering juga disebut model garis lurus (*A First-Order (Straight-Line) Model*). Bentuk umum Model Regresi Linear Sederhana diberikan oleh persamaan berikut.

$$y = \beta_0 + \beta_1 x + \varepsilon \qquad (6.1.2)$$

di mana

y = Variabel respon

x = Variabel prediktor

$E(y) = \beta_0 + \beta_1 x$ – Komponen deterministik

ε = Komponen galat acak

β_0 = Konstanta model regresi atau titik potong (*intercept*) garis regresi dengan sumbu-y (Nilai y di mana garis regresi memotong sumbu-y).

β_1 = Koefisien peubah prediktor x atau gradien garis regresi, yaitu besarnya kenaikan atau penurunan nilai E(y) untuk setiap kenaikan satu satuan x

Persamaan (6.1.2) adalah model dari yang kita yakini. Kita mulai dengan mengasumsikan bahwa model itu benar; namun nanti kita harus menyelidiki apakah memang demikian halnya. Dalam banyak aspek Statistika, kita sering harus mengasumsikan suatu model matematis untuk dapat mebuat kemajuan. Model ini harus diperiksa secara kritis pada suatu tahap, dan harus diubah jika ternyata tidak benar.

β_0 dan β_1 disebut parameter model regresi. Pemahaman model linear atau nonlinear didasarkan pada kelinearan atau kenonlinearan dari bentuk parameter di dalam model. Sedangkan pangkat tertinggi dari peubah peramal dalam model disebut orde model tersebut. Misalnya,

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon \qquad (6.1.3)$$

adalah suatu model regresi linear (dalam β) dan orde kedua (dalam x).

Selanjutnya, β_0 , β_1 dan ε tidak diketahui nilainya dalam Pers.(1), dan memang ε sangat sukar diketahui sebab nialinya berubah-ubah untuk setiap amatan y. Akan tetapi β_0 dan β_1 selalu tetap dan, meskipun kita tidak mungkin mengetahui berapa persis nilainya tanpa memeriksa semua kemungkinan pasangan y dan x, kita dapat menggunakan informasi di dalam data untuk menghasilkan dugaan (penaksir) bagi β_0 dan β_1 yaitu $\hat{\beta}_0$ dan $\hat{\beta}_1$. Jadi, kita dapat menuliskan estimasi untuk y dalam Persamaan (6.1.4) sebagai

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \qquad (6.1.4)$$

Persamaan merupakan persamaan prediksi bagi y; subsitusi untuk suatu nilai x akan menghasilkan prediksi bagi E(y) atau rata-rata populasi y pada nilai x tertentu.

Pendugaan untuk β_0 dan β_1 dilakukan dengan metode kuadrat terkecil (menggunakan Kalkulus) sehingga model regresi sering juga disebut model kuadrat terkecil dan garis regresi disebut juga garis kuadrat terkecil. Misalkan kita memiliki pasangan pengamatan (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) seperti yang diberikan pada Tabel 6.1.1.

Tabel 6.1.1 Pengamatan RLS

x_i	x_1	x_2	....	x_n
y_i	y_1	y_2	....	y_n

Menurut Persamaan (6.1.2) : $y = \beta_0 + \beta_1 x + \varepsilon$, maka kita dapat menuliskan

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad \forall\ i = 1, 2, \ldots, n \tag{6.1.5}$$

sehingga diperoleh :

$y_1 = \beta_0 + \beta_1 x_1 + \varepsilon_1$

$y_2 = \beta_0 + \beta_1 x_2 + \varepsilon_2$

……..

$y_n = \beta_0 + \beta_1 x_n + \varepsilon_n$

Untuk i = 1, 2, ..., n, didefenisikan galat ke-i :

$$\varepsilon_i = y_i - \beta_0 - \beta_1 x_i \quad (\text{dari : } y_i = \beta_0 + \beta_1 x_i + \varepsilon_i) \tag{6.1.6}$$

dan kuadrat galat adalah :

$$\varepsilon_i^2 = (y_i - \beta_0 - \beta_1 x_i)^2 \tag{6.1.7}$$

Selanjutnya didefinisikan Jumlah Kuadrat semua Galat (JKG) (*Sum of Squares of residuals/errors* = SSE) yaitu,

$$\text{JKG} = \sum_{i=1}^{n} [y_i - \beta_0 - \beta_1 x_i]^2 \tag{6.1.8}$$

yang merupakan fungsi 2-variabel β_0 dan β_1. Dapat dibuktikan bahwa dugaan/estimator untuk β_0 dan β_1 dengan cara Kalkulus ini akan memberikan nilai minimum untuk JKG. Itulah sebabnya penaksir untuk β_0 dan β_1 ini disebut penaksir kuadrat terkecil.

Penentuan penaksir untuk β_0 dan β_1 (seperti mencari titik kritis fungsi 2-variabel) :

$$\frac{d}{d\beta_0}[\text{SSE}] = \frac{d}{d\beta_0}\left\{\sum_{i=1}^{n} [y_i - \beta_0 - \beta_1 x_i]^2\right\} = 0 \text{ , dan } \beta_1 \text{ dianggap konstanta.} \tag{6.1.9}$$

$$\frac{d}{d\beta_1}[\text{SSE}] = \frac{d}{d\beta_1}\left\{\sum_{i=1}^{n} [y_i - \beta_0 - \beta_1 x_i]^2\right\} = 0 \text{ , dan } \beta_0 \text{ dianggap konstanta.} \tag{6.1.10}$$

Konsep ukuran ketepatan Model Regresi

Konsep ukuran ketepatan Model Regresi dimulai dari identitas yang diberikan dalam Persamaan (6.1.11) berikut :

$$y_i - \hat{y}_i = (y_i - \bar{y}) - (\hat{y}_i - \bar{y}) \qquad (6.1.11)$$

Simpangan prediksi/ramalan ($\hat{y}_i$) terhadap rata-rata ($\bar{y}$)

Simpangan amatan (y_i) terhadap rata-rata ($\bar{y}$)

Simpangan amatan (y_i) terhadap prediksi/ramalan $\hat{y}_i$ (galat/*error*)

Atau,

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}) \qquad (6.1.12)$$

Jika kedua ruas kiri dan kanan Persamaan (6.1.12) dikuadratkan dan kemudian kita jumlahkan dari i = 1, 2, ..., n maka kita diperoleh:

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n} (y_i - \hat{y})^2 + 2\sum_{i=1}^{n} (\hat{y}_i - \bar{y})(y_i - \hat{y}) \qquad (6.1.13)$$

$= 0$

JK simpangan antara amatan ke-i dengan nilai ramalannya(JKGalat = SSE)

JK simpangan antara nilai ramalan amatan ke-i dengan nilai rata-rata ($JK_{karenaregresi}$=JKR=SSR)

JK simpangan antara amatan ke-i dengan nilai rata-rata (JK di sekitar nilai tengah / JK terkoreksi=JKTotal=SST)

Atau disebut juga sebagai:

$$JK_{sekitar\ rataan} = JK_{karena\ regresi} + JK_{sekitar\ regresi} \qquad (6.1.14)$$

Ini menunjukkan bahwa, di antara variasi y di sekitar nilai rata-ratanya, sebagian variasi itu berasal dari garis regresi dan sebagian lainnya merupakan akibat dari kenyataan bahwa amatan-amatan itu tidak seluruhnya terletak pada garis regresi – seandainya seluruh amatan terletak pada garis regresi, maka jumlah kuadrat sekitar garis regresi akan sama dengan nol (*error* = 0). Dari pembahasan ini menjadi jelaslah bahwa salah satu cara untuk mengevaluasi kebaikan Model Regresi sebagai prediktor adalah menghitung berapa banyak JK sekitar nilai rata-rata terurai menjadi JK karena regresi dan JK sekitar regresi. Tentu sja diharapkan JK karena regresi jauh lebih besar daripada JK sekitar regresi, atau dengan kata lain kalau rasio atau nisbah

$$\frac{JK_{\text{karena regresi}}}{JK_{\text{sekitar rataan}}}$$

tidak terlalu jauh dari satu. Rasio ini disebut koefisien determinasi dan notasinya adalah R^2. Jadi,

$$R^2 = \frac{JK_{\text{karena regresi}}}{JK_{\text{sekitar rataan}}} = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \qquad (6.1.15)$$

di mana:

$$JK_{\text{karena regresi}} = \hat{\beta}_1^2 \sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad (6.1.16)$$

Jadi, R^2 mengukur proporsi variasi total di sekitar nilai tengah $\bar{y}$ yang dapat dijelaskan oleh regresi tersebut. Ukuran ini sering diucapkan sebagai persentase dengan mengalikannya dengan 100. R^2 dapat mengambil nilai maksimum 1 (atau 100%) bila semua nilai x berbeda.

Seperti telah dibahas sebelumnya, setiap jumlah kuadrat (JK) (distribusinya khi-kuadrat) memiliki apa yang disebut derajat bebas (*degrees of freedom* = df). Bilangan ini menunjukkan berapa banyak informasi yang bebas di antara n amatan y_1, y_2, ..., y_n yang dibutuhkan untuk mendapatkan JK itu. Misalnya, JK sekitar nilai rata-rata (JKT) membutuhkan (n–1) informasi yang bebas (sebab di antara bilangan-bilangan $y_1 - \bar{y}, y_2 - \bar{y}, ..., y_n - \bar{y}$, hanya (n – 1) di antaranya yang bebas karena jumlah semua n bilangan itu sama dengan nol). Kita dapat menghitung JK karena regresi (JKR) dari suatu fungsi y_1, y_2, ..., y_n yaitu $\hat{\beta}_1$, sehingga JK ini memiliki satu derajat bebas. Melalui pengurangan, JK sekitar regresi yang selanjutnya disebut JK sisaan/galat (JKG), memiliki (n – 2) derajat bebas. Ini mencerminkan kenyataan bahwa sisaan-sisaan tadi berasal dari sebuah model garis lurus yang mengharuskan pendugaan dua parameter. Secara umum, JKG mempunyai derajat bebas sebanyak (banyak amatan – banyaknya parameter yang diduga). Berdasarkan ini, kita dapat menyusun sebuah tabel yang disebut sebagai Tabel Analisis Variansi (*Analysis of Variance Table* atau ANOVA *Table*) yang diberikan pada Tabel 6.1.2 berikut ini.

Tabel 6.1.2 Tabel Analisis Variansi (ANOVA) RLS

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah
Karena Regresi	1	$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	KT_{Reg}
Sekitar Regresi (Sisaan)	n – 2	$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$s^2 = \frac{JK}{n-2}$
Total, terkoreksi	n – 1	$\sum_{i=1}^{n}(y_i - \bar{y})^2$	

Asumsi dalam Analisis Regresi

Asumsi-asumsi yang berkaitan dengan distribusi peluang (*probability distribution*) dalam Analisis Regresi diberikan berikut ini. Tinjau kembali model yang diberikan dalam Persamaan (6.1.17).

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \ , \ i = 1, 2, ..., n \qquad (6.1.17)$$

1. Galat ke-i (ε_i) merupakan suatu peubah acak dengan rata-rata nol dan variansi σ^2 yang tidak diketahui, yaitu:

   $E(\varepsilon_i) = 0$, $V(\varepsilon_i) = \sigma^2 \ \forall \ i = 1, 2, ..., n.$

2. ε_i merupakan suatu peubah acak Normal, dengan rata-rata nol dan variansi σ^2, yang dapat ditulis,

   $\varepsilon_i \sim N(0, \sigma^2)$

3. ε_i dan ε_j tidak berkorelasi, $i \neq j$, yaitu: $cov(\varepsilon_i, \varepsilon_j) = 0$

   Hal ini mengakibatkan:

   $$E(y_i) = E[\beta_0 + \beta_1 x_i + \varepsilon_i] = E[\beta_0 + \beta_1 x_i] + E[\varepsilon_i] = \beta_0 + \beta_1 x_i \qquad (6.1.18)$$

   $V(y_i) = var[\beta_0 + \beta_1 x_i + \varepsilon_i]$

   $= var[\beta_0 + \beta_1 x_i] + var[\varepsilon_i] + 2cov[\beta_0 + \beta_1 x_i, \varepsilon_i] = \sigma^2 \rightarrow V(y_i) = \sigma^2$

   dan y_i dan y_j tidak berkorelasi, $i \neq j$.

4. Dari 2 dan 3 maka y juga merupakan suatu peubah acak Normal, yang dapat ditulis,

   $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$

Berdasarkan asumsi-asumsi di atas, maka kita dapat menyusun prosedur inferensi terhadap parameter-parameter Model Regresi dengan menggunakan Distribusi Normal. Untuk itu sama halnya kita harus menaksir parameter-parameter β_0 dan β_1, maka dalam Analisis Regresi juga sangat penting untuk menaksir nilai σ^2 (yang tidak diketahui). Penaksir tak bias untuk σ^2 dan σ dalam Model Regresi Linear Sederhana diberikan dalam Persamaan (6.1.19).

$$S^2 = \frac{JKG}{n-2} \ , \ S = \sqrt{S^2} \qquad (6.1.19)$$

Inferensi parameter β_0 dan β_1 dalam Model Regresi Linear Berganda.

a. Selang kepercayaan $(1-\alpha)100\%$ untuk β_i , i = 0,1 :

$$\hat{\beta}_i \pm (t_{\alpha/2}) \ s_{\hat{\beta}_i}$$

dengan $t_{\alpha/2}$ mempunyai $\nu = n-2$.

b. Prosedur pengujian hipotesis parameter β_0 dan β_1 dalam Model Regresi Linear Berganda. diberikan pada Tabel 6.1.3 berikut.

Tabel 6.1.3 Uji hipotesis parameter β_i RLS

H_0	Uji Statistik	H_1	Daerah Kritis
$\beta_i = 0$	$t = \frac{\hat{\beta}_i}{s_{\hat{\beta}_i}}$	$\beta_i < 0$	$T < -t_\alpha$
		$\beta_i > 0$	$T > t_\alpha$
		$\beta_i \neq 0$	$T < -t_{\alpha/2}$ dan $t > z_{\alpha/2}$

Karena β_1 adalah besarnya kontribusi/pengaruh variabel x terhadap E[Y], maka hipotesis nol

$H_0 : \beta_1 = 0$

bermakna bahwa tidak ada pengaruh pengaruh variabel x terhadap E[Y]. Prosedur pengujian hipotesis untuk keputusan penolakan H_0 dilakukan dengan menggunakan Tabel 6.1.3 di atas.

Model Regresi Linear Berganda (*Multiple Regression Model*)

Model regresi yang banyak digunakan adalah Model Regresi Linear Berganda yang bentuk umunya diberikan pada Persamaan berikut.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_k + \varepsilon \qquad (6.1.20)$$

di mana

y adalah variabel respon

$x_1, x_2, \ldots, x_k$ adalah variabel-variabel prediktor

$E(y) = \beta_0 + \beta_1 x + \beta_2 x_2 + \ldots + \beta_p x_k$ adalah komponen deterministik

β_i adalah besarnya kontribusi/pengaruh variabel x_i terhadap E[Y], dengan menganggap beta lain konstan.

Bentuk Data dalam Model Regresi Linear Berganda (RLB) diberikan dalam Tabel berikut.

Tabel 6.1.4 Data Model RLB

Data Point	y	x_1	x_2	...	x_k
1	y_1	x_{11}	x_{21}	...	x_{k1}
2	y_2	x_{12}	x_{22}	...	x_{k2}
⋮	⋮	⋮	⋮	⋮	⋮
n	y_n	x_{1n}	x_{2n}	...	x_{kn}

Sama hal nya seperti pada Model Regresi Linear Sederhana, pendugaan parameter-parameter β_0, β_1, …., β_k juga dilakukan dengan metode kuadrat terkecil (menggunakan Kalkulus) sehingga diperoleh model penduga:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_k x_k \qquad (6.1.21)$$

yang meminimumkan Jumlah Kuadrat Galat (JKG) (Persamaan):

$$JKG = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (6.1.22)$$

Menurut Persamaan : $y = \beta_0 + \beta_1 x + \varepsilon$, kita dapat menuliskan :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i \qquad \forall\, i = 1, 2, \ldots, n \qquad (6.1.23)$$

sehingga diperoleh :

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \ldots + \beta_k x_{k1} + \varepsilon_1$$
$$y_2 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \ldots + \beta_k x_{k2} + \varepsilon_2$$
$$\ldots\ldots\ldots\ldots$$
$$y_n = \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} + \ldots + \beta_k x_{kn} + \varepsilon_n$$

Dan: $$JKG = \sum_{i=1}^{n} [y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - .... - \beta_k x_{ki}]^2 \qquad (6.1.24)$$

Penentuan penaksir β_i , i = 0, 1, 2,…, k.

$$\frac{d}{d\beta_i}\left\{\sum_{i=1}^{n}\left[y_i-\beta_0-\beta_1x_{1i}-\beta_2x_{2i}-....-\beta_kx_{ki}\right]^2\right\}=0\ ,\ i=0,1,2,\ldots,k \qquad (6.1.25)$$

Misal untuk menentukan penaksir β_0 diperoleh dari turunan parsial pertama JKG terhadap β_0 dengan menganggap β_I , I = 1, 2, …, k adalah konstan. Persamaan penaksir diberikan berikut.

$$\frac{d}{d\beta_0}\left\{\sum_{i=1}^{n}\left[y_i-\beta_0-\beta_1x_{1i}-\beta_2x_{2i}-....-\beta_kx_{ki}\right]^2\right\}=0 \qquad (6.1.26)$$

Asumsi-asumsi yang mendasari penggunaan Analisis Regresi Linear Berganda juga sama dengan asumsi-asumsi pada Analisis Regresi Linear Sederhana.

Inferensi parameter β_i dalam Model Regresi Linear Berganda.

a. Selang kepercayaan $(1-\alpha)100\%$ untuk β_i :

$$\hat{\beta}_i \pm (t_{\alpha/2})\ s_{\hat{\beta}_i}$$

dengan $t_{\alpha/2}$ mempunyai $\nu = n - (p+1)$.

b. Prosedur pengujian hipotesis parameter β_i dalam Model Regresi Linear Berganda. diberikan pada Tabel 6.1.5 berikut.

Tabel 6.1.5 Uji hipotesis parameter β_i RLB

H_0	Uji Statistik	H_1	Daerah Kritis
$\beta_i = 0$	$t = \frac{\hat{\beta}_i}{s_{\hat{\beta}_i}}$	$\beta_i < 0$	$T < -t_\alpha$
		$\beta_i > 0$	$T > t_\alpha$
		$\beta_i \neq 0$	$T < -t_{\alpha/2}$ dan $t > z_{\alpha/2}$

Karena β_i adalah besarnya kontribusi/pengaruh variabel x_i terhadap E[Y], maka hipotesis nol

$H_0 : \beta_i = 0$

bermakna bahwa tidak ada pengaruh pengaruh variabel x_i terhadap E[Y], dengan menganggap beta lain konstan. Prosedur pengujian hipotesis untuk keputusan penolakan H_0 dilakukan dengan menggunakan Tabel 6.1.5 di atas.

6.2. Artikel Ilmiah Penulis Menggunakan Analisis Regresi

6.2.1. Artikel Publikasi *Proceedings of the International Conference on Industrial Engineering and Operations Management* Monterrey, Mexico, November 3-5, 2021

Dalam artikel dengan judul *Statistical Analysis of the Relationship between the Sintering Process and Magnetic Energy on the Improvement of the Magnetic Properties of the Permanent Magnet* $BaFe_{12}O_{19}$. (Perak Samosir dkk, 2021) yang dipublikasikan pada *Proceedings of the International Conference on Industrial Engineering and Operations Management* Monterrey, Mexico, November 3-5, 2021, disampaikan bahwa magnet permanen merupakan salah satu bahan pendukung komponen dalam peralatan elektronik. Secara khusus pada generator putaran rendah biasanya digunakan magnet NdFeB. Alternatif lain yang dapat digunakan adalah material multiferoik berbasis ferit $BaFe_{12}O_{19}$ (barium heksaferit). $BaFe_{12}O_{19}$ dikenal memiliki sifat magnetik yang baik dan banyak digunakan dalam perangkat magnetik untuk menyimpan energi dalam medan magnet statik. Hal ini dapat dilihat dari nilai energi magnetik material tersebut, yang akan menentukan kualitas magnet dan selanjutnya akan mempengaruhi daya listrik (Suastiyanti, D. dkk, 2019).

Andreas, K. dkk (2018) telah melakukan penelitian untuk memperoleh sifat magnetik tertinggi pada $BaFe_{12}O_{19}$ melalui variasi perlakuan proses sintering. Variabel proses sintering penelitian adalah temperatur proses sintering (750°C, 800°C dan 850°C) dan lama proses sintering (8 jam, 10 jam dan 12 jam). Sifat magnetik yang diteliti adalah sifat remanen, sifat koersivitas dan sifat saturasi. Penelitian dilakukan dengan menggunakan metode *Thermo Gravimetric Analyzer*/ *Differential Thermal Analysis* (TGA/DTA). Samosir, P. dan Simamora, D. B. S (2020) menganalisis data Andreas, K. dkk (2018) di atas dengan menggunakan Analisis Regresi untuk melihat hubungan antara proses sintering dengan sifat magnetik material multiferoik berbasis ferit dan dari hasil Analisis Statistika diperoleh bahwa ada hubungan yang signifikan antara proses sintering dengan sifat remanen, koersivitas dan saturasi.

Selanjutnya Samosir, P. dkk (2021) menganalisis hubungan antara proses sintering dengan energi magnetik pada peningkatan sifat magnetik magnet permanen $BaFe_{12}O_{19}$ juga dengan menggunakan Analisis Regresi. Variabel-variabel prediktor penelitian adalah variabel proses sintering yang terdiri dari temperatur sinter (°C) dan waktu yang digunakan untuk proses sintering (jam). Sedangkan variabel respon penelitian adalah energi magnetik. Data nilai Energi Magnetik (EM) diberikan pada Tabel 6.2.1.1 berikut.

Tabel 6.2.1.1 Nilai Energi Magnetik

Temperatur (°C)	Waktu		
	8 jam	10 jam	12 jam
750	19,5358	20,01	23,0317
800	22,8765	45,6192	72,8178
850	36,1669	62,8524	79,7791

Dengan menggunakan Analisis Regresi kita dapat melihat besarnya pengaruh temperatur dan waktu sintering terhadap nilai rata-rata energi magnetik. Model yang dihasilkan pun dapat digunakan untuk melakukan prediksi nilai rata-rata energi magnetik jika proses sintering dilakukan pada temperatur dan waktu tertentu.

Model Regresi Linear Berganda untuk data pada Tabel 6.2.1.1 adalah

$$y = E(y) + \varepsilon \qquad (6.2.1.1)$$

di mana komponen deterministik diberikan dalam Persamaan 2 berikut.

$$E(y) = \beta_0 + \beta_1 x + \beta_2 x_2 \qquad (6.2.1.2)$$

dimana y adalah nilai Energi Magnetik (EM) dalam T.kA/m, x_1 adalah temperatur proses sintering (^oC) dan x_2 adalah waktu proses sintering (jam).

Dalam model pada Persamaan (6.2.2) di atas terdapat tiga buah parameter yaitu β_0, β_1 dan β_2. Agar dapat mengestimasi nilai y maka langkah selanjutnya adalah menentukan nilai ketiga parameter β_0, β_1 dan β_2 dengan menggunakan data pada Tabel 6.2.1.1. Jika nilainya sudah diestimasi maka kita dapat menentukan nilai $\hat{y}$, yang diberikan dalam Persamaan 6.2.1.3 berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 \qquad (6.2.1.3)$$

Hipotesis nol (H_0) adalah tidak ada hubungan antara proses sintering dengan energi magnetik pada peningkatan sifat magnetik magnet permanen $BaFe_{12}O_{19}$. Hipotesis alternatif (H_1) adalah ada hubungan antara proses sintering dengan energi magnetik pada peningkatan sifat magnetik magnet permanen $BaFe_{12}O_{19}$. Hipotesis ini ditulis sebagai:

$H_0 : \beta_{1,0} = \beta_{2,0} = 0$ lawan alternatif H_1 : minimal salah satu beta $\neq 0$

Kriteria penolakan H_0 dapat dilihat dari nilai-p pada Tabel Koefisien pada output pengolahan data. Jika nilai-p (p-*value*) lebih kecil dari kriteria taraf signifikansi $\alpha = 0.05$, maka H_0 ditolak, dan uji dikatakan signifikan, yang bermakna ada hubungan yang signifikan antara proses sintering dengan energi magnetik pada peningkatan sifat magnetik magnet permanen $BaFe_{12}O_{19}$. Jika p-value lebih kecil dari 0.01, uji dikatakan sangat signifikan, yang bermakna ada

hubungan yang sangat signifikan antara proses sintering dengan energi magnetik pada peningkatan sifat magnetik magnet permanen $BaFe_{12}O_{19}$. Jika p-value lebih besar dari kriteria α = 0.05, maka H_0 diterima bahwa tidak ada hubungan antara proses sintering dengan energi magnetik pada peningkatan sifat magnetik magnet permanen $BaFe_{12}O_{19}$.

Dari hasil pengolahan data diperoleh persamaan regresi sebagai berikut:

$$\widehat{EM} = -348.3 + 0.3874 \text{ Temperatur} + 8.09 \text{ Waktu} \qquad (6.2.1.4)$$

Hasil uji hipotesis untuk parameter model regresi linear berganda diberikan pada Tabel 6.2.1.2 berikut.

Tabel 6.2.1.2 Hasil uji hipotesis parameter model energi magnetik

Term	Koefisien	Nilai-t	Nilai-p
Constant	–348.3	–4.55	0.004
Temperatur	0.3874	4.25	0.005
Waktu	8.09	3.55	0.012

Berdasarkan Tabel 6.2.1.2 di atas diperoleh bahwa nilai-p untuk $\hat{\beta}_1$ adalah 0,005 dan untuk $\hat{\beta}_2$ adalah 0,012. Kedua nilai-p ini lebih kecil dari α = 0,05. Dengan demikian untuk hipotesis,

$H_0 : \beta_{1,0} = \beta_{2,0} = 0$ lawan alternatif H_1 : minimal salah satu beta $\neq 0$

H_0 ditolak yang bermakna ada hubungan yang sangat signifikan antara proses sintering dengan energi magnetik pada peningkatan sifat magnetik magnet permanen $BaFe_{12}O_{19}$.

Dari Tabel 6.2.1.2 di atas juga diperoleh bahwa:

$$\hat{\beta}_1 = 0{,}3874 \text{ dan } \hat{\beta}_2 = 8{,}09$$

Hal ini berarti untuk setiap kenaikan 1°C temperatur sintering, maka nilai rata-rata Energi Magnetik ($\widehat{EM}$) akan naik sebesar 0.3874 T.kA/m dan untuk setiap kenaikan 1jam waktu sintering, nilai rata-rata Energi Magnetik ($\widehat{EM}$) akan naik sebesar 8.09 T.kA/m.

Dengan demikian kita dapat menggunakan model untuk membuat suatu selang kepercayaan terhadap nilai rata-rata energi mekanik (EM) untuk nilai tertentu temperatur dan waktu proses sintering seperti berikut ini.

1. Temperatur 750°C dan waktu 12jam.

Tabel 6.2.1.3 Prediksi Temperatur 750°C dan waktu 12jam

Fit	Selang Kepercayaan 95%
39,3258	(21,0970 – 57,5546)

2. Temperatur 850°C dan waktu 12jam.

Tabel 6.2.1.4 Prediksi Temperatur 850°C dan waktu 12jam

Fit	Selang Kepercayaan 95%
78,0661	(59,8373 – 96,2949)

3. Temperatur 900°C dan waktu 14jam.

Tabel 6.2.1.5 Prediksi Temperatur 900°C dan waktu 14jam

Fit	Selang Kepercayaan 95%	
113,611	(80,7488 – 146,474)	X

X denotes an unusual point relative to predictor levels used to fit the model.

Dalam prediksi yang ketiga, data yang dipakai adalah nilai temperatur dan waktu proses sintering bukan data yang dipakai untuk menyusun model. Hal ini terlihat dalam *output Prediction* (yang di*highlight*). Dengan menggunakan model ini kita dapat memprediksi bahwa jika diambil temperatur 900°C dan waktu = 14 jam, diperoleh nilai rata-rata Energi Magnetik 113,611 T.kA/m.

6.2.2. Artikel Publikasi *International Journal of Innovation, Creativity and Change*.

www.ijicc.net Volume 14, Issue 10, 2020

Samosir, P & Rajagukguk, W (2020) telah melakukan penelitian dengan judul *Infrastructure Development and Economic Growth in Indonesia: A Province Panel Data Analysis*. Artikel hasil penelitian telah diterbitkan pada *International Journal of Innovation, Creativity and Change*. www.ijicc.net Volume 14, Issue 10, 2020. Dalam artikel ini disampaikan bahwa pembangunan infrastruktur dipandang sebagai salah satu faktor pendorong pembangunan. Pada sisi lain, pembangunan infrastruktur yang tidak berimbang berdampak pada kesenjangan pertumbuhan ekonomi (Rajagukuguk, W, 2016). Penelitian ini menyelidiki pengaruh pembangunan infrastruktur dengan pertumbuhan ekonomi dan disparitas pertumbuhan ekonomi antar Provinsi di Indonesia yang terdiri atas 33 Provinsi. Disparitas dianalisis antar provinsi. Data yang digunakan adalah data panel Provinsi tahun 2001 sampai dengan tahun 2012. Data

diperoleh dari *The Indonesia Database for Policy and Economic Research* (INDO-DAPOER) *World Bank*.

Dalam penelitian ini Analisis Statistika dilakukan dengan menggunakan Analisis Regresi Data Panel yaitu *Fixed Effect Model* (FEM), yang merupakan salah satu modifikasi Analisis Regresi. Variabel respon adalah angka pertumbuhan Produk Domestik Bruto (PDB) per kapita tanpa harga konstan (*constant prices*) minyak dan gas (dalam juta rupiah), dan variabel prediktor adalah *Household Access to Electricity* (Akses Listrik Rumah Tangga = ALRT), Total (dalam % dari total rumah tangga), Angka Melek Huruf (AMH) penduduk usia 15 tahun ke atas (dalam % dari total populasi), dan Persentase Kelahiran (PK) yang ditangani oleh tenaga medis (dalam % dari total kelahiran).

Dalam penyusunan model dengan menggunakan Analisis Regresi bentuk data akan menentukan model regresi yang dapat digunakan. Dalam kasus data yang melibatkan variabel-variabel ekonomi biasanya data berbentuk data *cross section* atau data *time series*.

Data *cross section* (data silang) adalah data yang memiliki objek yang banyak pada tahun yang sama atau data yang dikumpulkan dalam satu waktu terhadap banyak objek. Pengertian objek di sini bisa macam-macam dan berupa banyak hal seperti misalnya individu/orang, perusahaan, bank, daerah (kabupaten dan kota), dan bahkan negara. Berbeda dengan *cross section*, data *time series* mengandung unsur waktu. konsep dari data *time series* (data runtut waktu) adalah data yang memiliki runtun waktu yang lebih dari satu tahun pada satu objek atau data yang dikumpulkan dari waktu ke waktu terhadap satu individu/objek. Misalnya peneliti ingin melakukan analisis trend konsumsi listrik dari tahun 1990 sampai dengan 2017. Data yang diambil biasanya hanya satu kelompok, yakni konsumsi listrik nasional namun memiliki *range* waktu yang panjang dari periode tahun tersebut. Tentu jumlah tahun akan sangat mempengaruhi pola data mengingat dalam berjalannya periode terjadi kondisi yang berbeda-beda, seperti tingkat inflasi, tingkat suku bunga, nilai mata uang, dan lain- lain. Metode yang digunakan untuk menganalisis kedua data ini berbeda.

Data pada penelitian ini adalah data panel. Data panel adalah kombinasi antara data *cross section* dan data *time series*. Analisis Regresi yang digunakan untuk mengolah data panel disebut Analisis Regresi Data Panel. Terdapat beberapa metode yang biasa digunakan dalam mengestimasi model regresi data panel, yaitu *pendekatan Common Effect Model (CEM), Fixed Effect Model (FEM*) dan *Random Effect Model (REM*) (Jeffrey M Wooldridge, 2010).

Dalam penelitian ini digunakan pendekatan FEM untuk menganalisis hubungan antara angka pertumbuhan Produk Domestik Bruto (PDB) per kapita tanpa harga konstan (*constant prices*) minyak dan gas, dengan variabel-variabel *Household Access to Electricity* (Akses Listrik Rumah Tangga). pada ke-33 provonsi di Indonesia.

Fixed Effect Model (FEM) untuk data panel dalam penelitian ini (Jeffrey M Wooldridge, 2010) adalah:

$$Y_{it} = \beta_{0,i} + \beta_1 X_{1,it} + \ldots + \beta_k X_{k,it} + u_{ii} \quad (6.2.2.1)$$

di mana :

i = 1,...,n, t = 1,...T,

n : banyak unit *cross sction*, T : banyak unit *time series*

β_0 : *Intercept* model regresi

Y_{it} : Nilai variabel respon *cross section* ke-*i time series* ke-*t*,

X_{it} : Nilai variabel bebas ke-*k* untuk *cross section* ke-*i* tahun ke-*t*,

β_i : parameter dari variabel prediktor ke-k

u_{ii} : komponen galat/gangguan.

Model di atas dapat dituliskan kembali sebagai berikut :

$$Y_{it} = \beta_0 + \beta_1.X_{1,it} + ... + \beta_2.X_{2.it} + \gamma_2.D2_i + ... + \gamma_2.Dn_i + \mu_i \quad (6.2.2.2)$$

dimana $D2_i$, …, Dn_i menyatakan variabel boneka (*dummy variable*) untuk provinsi. Untuk menaksir parameter pada model regresi data panel dengan *intercept* bervariasi dan *slope* konstan, pemodelan efek tetap komponen dua arah secara umum dilakukan dengan *Least Square Dummy Variable* (LSDV).

Hipotesis nol adalah tidak ada hubungan antara angka pertumbuhan Produk Domestik Bruto (PDB) per kapita tanpa harga konstan (*constant prices*) minyak dan gas (dalam juta rupiah) dengan variabel prediktor adalah *Household Access to Electricity* (Akses Listrik Rumah Tangga = ALRT), Total (dalam % dari total rumah tangga), Angka Melek Huruf (AMH) penduduk usia 15 tahun ke atas (dalam % dari total populasi), dan Persentase Kelahiran (PK) yang ditangani oleh tenaga medis (dalam % dari total kelahiran). Hipotesis alternatifnya adalah ada hubungan antara angka pertumbuhan Produk Domestik Bruto (PDB) per kapita tanpa harga konstan (*constant prices*) minyak dan gas (dalam juta rupiah) dengan variabel prediktor adalah *Household Access to Electricity* (Akses Listrik Rumah Tangga = ALRT), Total (dalam % dari

total rumah tangga), Angka Melek Huruf (AMH) penduduk usia 15 tahun ke atas (dalam % dari total populasi), dan Persentase Kelahiran (PK) yang ditangani oleh tenaga medis (dalam % dari total kelahiran).

Uji statistik yang digunakan adalah Uji Hausmann dengan kriteria ketepatan *Fixed Effect Model (FEM*) menggunakan nilai koefisien determinasi R^2 ($0\% \leq R^2 \leq 100\%$) dan nilai-p amatan ($0 \leq$ nilai-p $< \alpha$=0,05). Semakin akurat model menggambarkan hubungan antara variabel terikat dengan variabel bebas, maka semakin tinggi nilai R^2 dan nilai-p < 0,05.

Statistik deskriptif variabel-variabel dalam penelitian yaitu: jumlah observasi, rata-rata, standar deviasi, nilai minimum, dan maksimum diberikan pada pada Tabel 6.2.2.1.

Tabel 6.2.2.1 Statistik deskriptif variabel-variabel *Panel Data Analysis*

Variabel	Obs.	Mean	Std.Deviasi	Min	Maks
PDB per Kapita	396	7.312207	6.560	1.7589	43.1954
ALRT	396	81.80682	15.569	36.2	100
AMH	396	92.04975	6.183	64.1	99.3
PK	396	69.72475	16.943	27.9	99.3

Sumber: INDODAPOER, 2019. Dihitung peneliti.

Hasil analisis data menggunakan Uji Hausmann terhadap parameter *Fixed Effect Model* tersaji pada Tabel 6.2.2.2.

Tabel 6.2.2.2 Hasil uji Hausmann terhadap parameter *Fixed Effect Model*

Variabel	Koefisien	Nilai-t	Nilai-p
Konstanta	-1,559110	-5,372301	0,0000
ALRT	0,010446	13,83107	0,0000
AMH_	0,025831	7,296687	0,0000
PK	0,001508	2,523708	0,0120

Variabel respon:
Angka Pertumbuhan PDB per kapita tanpa harga konstanminyak dan gas (juta rupiah)
Sumber, Indodapoer, 2019, Dihitung peneliti.

Dari Tabel 6.2.2.2 di atas diperoleh bahwa nilai-p untuk semua parameter lebih kecil dari kriteria α = 0,05. Oleh karena hipotesis nol ditolak dan ada hubungan antara angka pertumbuhan Produk Domestik Bruto (PDB) per kapita tanpa harga konstan (*constant prices*) minyak dan gas (dalam juta rupiah) dengan variabel prediktor adalah *Household Access to Electricity* (Akses Listrik Rumah Tangga = ALRT), Total (dalam % dari total rumah tangga), Angka Melek Huruf (AMH) penduduk usia 15 tahun ke atas (dalam % dari total populasi), dan Persentase Kelahiran (PK) yang ditangani oleh tenaga medis (dalam % dari total kelahiran).

Secara nasional pertumbuhan ekonomi dapat dituliskan dalam persamaan berikut:

Pertumbuhan Ekonomi =

$$-1.559110 + 0.010446.\ ALRT + 0.025831\ AMH + 0.001508PK \quad (6.2.2.3)$$

Setiap peningkatan satu persen ALRT meningkatkan rata-rata angka pertumbuhan GDP per kapita sebesar 0.010446. Setiap kenaikan 1% AMH akan menaikkan rata-rata angka pertumbuhan GDP per kapita sebesar 0.025831. Selanjutnya setiap kenaikan 1% PK akan meningkatkan rata-rata angka pertumbuhan GDP Per kapita sebesar 0.001508. Dihasilkan bahwa semua variabel bebas berdampak positif pada pertumbuhan ekonomi. Hasil analisis juga menunjukkan bahwa terdapat disparitas pertumbuhan ekonomi akibat perbedaan konstanta pertumbuhan ekonomi antar provinsi di Indonesia.

6.2.3. Artikel Publikasi *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management*, Detroit, Michigan, USA, *August* 10 - 14, 2020

a. *Dark Side of Economic Growth: A Case Study of the Relationship between Economic Growth and Suicide Mortality*

Wilson, R. dkk (2020) telah melakukan penelitian dengan judul *Dark Side of Economic Growth: A Case Study of the Relationship between Economic Growth and Suicide Mortality* yang telah dipresentasikan dan dipublikasikan pada *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management*, Detroit, Michigan, USA, *August* 10 - 14, 2020. Dalam artikel ini disampaikan bahwa Pertumbuhan ekonomi sering dikaitkan dengan peningkatan kesejahteraan. Namun, di balik pertumbuhan ekonomi ada sisi gelap, antara lain kematian akibat bunuh diri.

Dalam penelitian ini dilakukan analisis untuk mengetahui hubungan antara pertumbuhan ekonomi dengan kematian akibat bunuh diri. Penelitian ini menggunakan data dari 180 negara di dunia pada tahun 2005, 2010, 2015, dan 2016. Oleh karena itu jumlah data penelitian ini adalah 720 untuk 4 tahun. Variabel respon adalah angka kematian bunuh diri per 100.000 penduduk (*Suicide mortality rate per 100,000 population*), angka kematian bunuh diri perempuan per 100.000 penduduk wanita (*Female suicide mortality rate per 100,000 female population*), dan angka kematian bunuh diri laki-laki per 100.000 penduduk laki-laki (*Male suicide mortality rate per 100,000 male population*). Sementara itu, variabel prediktor adalah pertumbuhan Produk Domestik Bruto (PDB) per kapita (tahunan %) (*Gross Domestic Product (GDP) per capita growth (annual %),*), tingkat deteksi kasus tuberkulosis (%, semua bentuk)

(*Tuberculosis case detection rate (%, all forms)*), dan total konsumsi alkohol per kapita (liter alkohol murni, perkiraan proyeksi, usia > 15 tahun) (*Total alcohol consumption per capita (liters of pure alcohol, projected estimates, 15+ years of age)*). Metode yang digunakan adalah Analisis Regresi Linier Berganda Multivariat.

Hipotesis nol penelitian adalah tidak ada hubungan antara ketiga variabel respon yaitu angka kematian bunuh diri (per 100.000 penduduk), angka kematian bunuh diri perempuan (per 100.000 penduduk wanita), dan angka kematian bunuh diri laki-laki (per 100.000 penduduk laki-laki) dengan variabel prediktor adalah pertumbuhan Produk Domestik Bruto (PDB) per kapita (tahunan %), tingkat deteksi kasus tuberkulosis (%, semua bentuk), dan total konsumsi alkohol per kapita (liter alkohol murni, perkiraan proyeksi, usia > 15 tahun). Hipotesis alternatif tentu saja ada hubungan antara ketiga variabel respon yaitu angka kematian bunuh diri (per 100.000 penduduk), angka kematian bunuh diri perempuan (per 100.000 penduduk wanita), dan angka kematian bunuh diri laki-laki (per 100.000 penduduk laki-laki) dengan variabel prediktor adalah pertumbuhan Produk Domestik Bruto (PDB) per kapita (tahunan %), tingkat deteksi kasus tuberkulosis (%, semua bentuk), dan total konsumsi alkohol per kapita (liter alkohol murni, perkiraan proyeksi, usia > 15 tahun).

Pada Tabel 6.2.3.1 berikut ini diberikan hasil pengolahan terhadap data dari 180 negara di dunia pada tahun 2005, 2010, 2015, dan 2016.

Dari Tabel 6.2.3.1 di atas diperoleh semua nilai-p untuk variabel prediktor lebih kecil dari kriteria $\alpha = 0{,}05$. Oleh karenanya hipotesis nol ditolak yang bermakana bahwa ada hubungan antara ketiga variabel respon yaitu angka kematian bunuh diri (per 100.000 penduduk), angka kematian bunuh diri perempuan (per 100.000 penduduk wanita), dan angka kematian bunuh diri laki-laki (per 100.000 penduduk laki-laki) dengan variabel prediktor adalah pertumbuhan Produk Domestik Bruto (PDB) per kapita (tahunan %), tingkat deteksi kasus tuberkulosis (%, semua bentuk), dan total konsumsi alkohol per kapita (liter alkohol murni, perkiraan proyeksi, usia > 15 tahun).

Tabel 6.2.3.1 Hasil pengolahan data *Dark Side of Economic Growth*

Variables, Coefficient, Standard Error, t, P-value, and 95% Confidence Interval

Variables	*Coefficient*	*Std. Err.*	*t*	*P-value*	*95% Conf. Interval*	
Suicide mortality rate per 100,000 population						
GDP per capita growth (annual %)	0.1680977	0.055836	3.01	0.003	0.0584766	0.27772
Tuberculosis case detection rate (%, all forms)	0.0770793	0.012074	6.38	0.000	0.0533769	0.100782
Total alcohol consumption per capita (liters of pure alcohol, projected estimates, 15+ years of age)	0.36022	0.012073	6.29	0.000	0.2478721	0.472568
Constant	2.710426	0.897907	3.02	0.003	0.9475809	4.47327
Female suicide mortality rate per 100,000 female population						
GDP per capita growth (annual %)	0.089784	0.030648	2.93	0.004	0.0296133	0.149955
Tuberculosis case detection rate (%, all forms)	0.0133171	0.006627	2.01	0.045	0.000307	0.026327
Total alcohol consumption per capita (liters of pure alcohol, projected estimates, 15+ years of age)	0.136786	0.031410	4.35	0.000	0.0751181	0.198453
Constant	3.540719	0.492859	7.18	0.000	2.573098	4.50834
Suicide mortality rate male per 100.000 population						
GDP per capita growth (annual %)	0.268161	0.093174	2.88	0.004	0.085234	0.451088
Tuberculosis case detection rate (%, all forms)	0.145511	0.020146	7.22	0.000	0.105958	0.185064
Total alcohol consumption per capita (liters of pure alcohol, projected estimates, 15+ years of age)	0.613820	0.09549	6.43	0.000	0.426342	0.801298
Constant	1.547238	1.498357	1.03	0.302	-1.39446	4.488935

Source: World Bank (2020) (Authors' calculation).

Hasil analisis data pada Tabel 6.2.3.1 di atas juga memperlihatkan semua nilai penaksir parameter variabel prediktor adalah positif. Hal ini bermakna bahwa pertumbuhan PDB per kapita, angka deteksi kasus tuberkulosis, dan total konsumsi alkohol per kapita memiliki hubungan positif dan signifikan dengan tiga angka kematian bunuh diri di dunia. Jadi temuan dari penelitian ini menegaskan bahwa ada sisi gelap dari pertumbuhan ekonomi. Pencapaian ekonomi yang lebih baik dapat menyebabkan tingkat kematian bunuh diri yang lebih tinggi sehingga negara-negara harus berhati-hati dan mengantisipasi tren yang meningkat dengan kebijakan yang bertujuan untuk menguranginya (Rajagukguk, W & Purba, J.T., 2012).

b. *E-Business Services Strategy with Financial Technology: Evidence from Indonesia*

Purba dkk (2020) telah melakukan penelitian dengan judul *E-Business Services Strategy with Financial Technology: Evidence from Indonesia*. Artikel hasil penelitian ini telah

dipresentasikan dan dipublikasikan pada *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management*, Detroit, Michigan, USA, August 10 - 14, 2020. Tujuan penelitian ini adalah untuk meneliti tentang bisnis teknologi keuangan (*financial technology*, selanjutnya disingkat *fintech*) dalam layanan bisnis elektronik yang ditawarkan oleh salah satu perusahaan *fintech* di Indonesia. Penelitian ini mencoba menganalisis keberadaan dan strategi layanan yang ditawarkan oleh perusahaan *fintech*. Hal ini penting karena dalam kehidupan sehari-hari sektor pembayaran digital semakin banyak yang menggunakan layanan *fintech* dan semakin berkembang di Indonesia. Ada dua pemain raksasa seperti GoPay dan OVO yang kini menguasai pasar di Indonesia. Kedua layanan ini disukai oleh masyarakat yang ingin mempermudah transaksi pembayaran (Tan, J. D. dkk. 2019). Mengingat perekonomian Indonesia telah meningkat dalam dua dekade terakhir, industri keuangan pun telah ikut bergerak karena kemajuan dan perkembangan teknologi yang cepat. Sektor keuangan di suatu negara memegang peranan penting dalam perekonomian dan terus mengalami perkembangan yang berkesinambungan sesuai dengan kebutuhan masyarakat.

Aplikasi Fintech yang disurvei dalam penelitian ini adalah Gopay yang dimiliki oleh PT. Gojek Indonesia yang berkantor pusat di Jakarta yang menjadi dompet digital yang paling banyak digunakan tahun 2020. Data penelitian merupakan data *cross section* hasil survey dengan responden adalah kalangan milenial wilayah Jakarta, Bogor, Depok, Tangerang dan Bekasi. Apalagi saat ini memasuki Revolusi Industri 4.0, kita perlu mengembangkan norma dan protokol umum untuk memastikan bahwa teknologi melayani umat manusia dan berkontribusi untuk masa depan yang lebih baik dan berkelanjutan (Cann, 2016). Berkaitan dengan laporan Cann, dimulainya *fintech* semakin maraknya perkembangan layanan *fintech* di Indonesia, sehingga keamanan data pribadi dituntut untuk terus ditingkatkan guna memberikan kontribusi gaya hidup yang lebih baik (Purba, 2016). Maka variabel respon dalam penelitian ini adalah preferensi pengguna (senang atau tidak senang menggunakan aplikasi GoPay *Fintech*), dan variabel prediktor adalah layanan *fintech* yaitu manfaat (*benefit*) dan keamanan (*security*) dalam penggunaan aplikasi GoPay *Fintech*.

Penelitian bertujuan untuk memperjelas seberapa baik dampak layanan *fintech*, manfaat dan keamanan terhadap preferensi mereka dalam bertransaksi secara *online*. Data dianalisis dengan menggunakan Analisis Regresi bersamaan dengan Analisis Regresi *Robust*. Analisis Regresi *Robust* merupakan salah satu modifikasi dari Analisis Regresi ketika asumsi Distribusi

Normal tidak dipenuhi. Biasanya hal ini terjadi akibat adanya data pencilan (*outlier*) sehingga menyebabkan bentuk distribusi yang nonnormal. Analisis Regresi *Robust* dapat mengatasi masalah ini. Oleh karenanya untuk data ini, pengolahan data dilakukan dengan menggunakan aplikasi Stata versi 15.

Hipotesis yang akan diuji adalah apakah responden senang menggunakan layanan yang diberikan oleh *fintech* khususnya yang merupakan kalangan milenial. Rumusan hipotesisnya diberikan sebagai berikut.

H_0 : Tidak ada hubungan antara preferensi pengguna dengan layanan GoPay *Fintech*

H_1 : Ada hubungan antara preferensi pengguna dengan layanan GoPay *Fintech*

Pada Tabel 6.2.3.2 berikut diberikan hasil pengolahan data.

Tabel 6.2.3.2 Hasil pengolahan data *Financial Technology*

Source	SS	df	MS	Number of obs	=	238
				F(2, 235)	=	196.05
Model	896.975517	2	448.487759	Prob > F	=	0.0000
Residual	537.595911	235	2.28764218	R-squared	=	0.6253
				Adj R-squared	=	0.6221
Total	1434.57143	237	6.053044	Root MSE	=	1.5125

CP	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Benefit	.5297602	.0395121	13.41	0.000	.451917	.6076034
Security	.2756795	.0441409	6.25	0.000	.1887171	.3626419
_cons	1.660268	.7953475	2.09	0.038	.0933455	3.22719

Dari Tabel di atas terlihat bahwa nilai-p untuk kedua dan variabel prediktor *benefit* dan *security* masing-masing adalah 0,000. Oleh karena itu hipotesis nol ditolak dan kedua variabel prediktor mempengaruhi preferensi pengguna GoPay *Fintech*. Tampaknya konsumen cenderung memilih untuk memperhatikan *benefit* dan *security* yang menurut mereka sesuai dengan keinginan mereka. Dengan pilihan ini setiap konsumen secara individual akan mencoba untuk mencerminkan identitas mereka sendiri melalui pilihan *benefit* dan *security*.

c. *Measurement Availability of Clean Water and ElementaryTeachers towards Income of all Districts and Cities in Indonesia*

Sidik Budiono dkk (2020) telah melakukan penelitian dengan judul *Measurement Availability of Clean Water and Elementary Teachers towards Income of all Districts and*

Cities in Indonesia yang telah dipresentasikan dan dipublikasikan pada *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management*, Detroit, Michigan, USA, August 10 - 14, 2020. Dalam penelitian ini disampaikan bahwa kinerja suatu daerah dapat tercermin dari pencapaian tingkat pendapatan di daerah tersebut. Dalam *Business Cycle Theory* begitu banyak faktor ekonomi dan non ekonomi yang mempengaruhi tingkat pendapatan. Kajian dalam penelitian ini ingin mengungkap betapa pentingnya ketersediaan air bersih dan guru SD dalam mendukung kesejahteraan ekonomi lokal di setiap kabupaten dan kota di Indonesia yang disurvei. Kehadiran guru sekolah dasar memiliki peran mendidik siswa dan membuat lingkungan masyarakat lebih sadar untuk meningkatkan kehidupan yang lebih baik. Air bersih dapat berperan dalam memenuhi kebutuhan pokok dan kegiatan secara lebih efisien serta dapat menentukan keputusan ekonomi. Penelitian ini menggunakan data cross section dengan 501 kabupaten dan kota di seluruh Indonesia.

Data air bersih adalah jumlah pelanggan rumah tangga pada perusahaan publik air bersih. Sarana air bersih di Indonesia dikelola oleh perusahaan negara atau pemerintah daerah yang berada di wilayah-wilayah tersebut. Jumlah guru SD yang melayani kegiatan mengajar di seluruh kabupaten dan kota di seluruh Indonesia. Guru berperan tidak hanya dalam mengajar tetapi juga memberikan nasehat-nasehat penting dalam bidang lingkungan sosialnya.

Asumsi peneliti ada dua hal utama yang memiliki kontribusi besar bagi kesejahteraan masyarakat Indonesia. Pertama, akses air bersih. Air bersih merupakan kebutuhan publik yang penting bagi masyarakat. Air bersih digunakan masyarakat untuk aktivitas sehari-hari sehingga segala aktivitas menjadi lebih hemat (Purba, J. T. & Budiono, S. 2019). Kedua, peran guru SD terhadap siswa dan lingkungan masyarakat di wilayah Indonesia. Guru-guru SD di Indonesia telah dilantik untuk ditempatkan dimanapun di Indonesia. Dengan harapan guru SD berperan aktif dan membuka pandangan siswa dan lingkungan sekitarnya untuk berpikir jauh ke depan. Guru SD merupakan salah satu agen perubahan yang krusial (Budiono, 2009). Maka variabel-variabel prediktor dalam penelitian ini adalah jumlah rumah tangga yang dapat mengakses air bersih di kabupaten dan kota (*The number of households that can access clean water in regencies and cities*), diberi label HH_W dan jumlah guru SD di kabupaten dan di kota (*Number of elementary school teachers in districts and cities*), diberi label EST. Sedangkan variabel respon adalah Produk Domestik Regional Bruto (PDRB) (*Gross Domestic Regional Product*), diberi

label GDRP. Data merupakan data *cross section* di 501 observasi (kabupaten dan kota) di Indonesia. Analisis data dilakukan dengan menggunakan Analisis Regresi Linear Berganda.

Hipotesis yang akan diuji adalah apakah ada pengaruh HH_W dan EST terhadap GDRP, sehingga rumusan hipotesisnya adalah sebagai berikut.

H_0 : Tidak ada hubungan antara HH_W dan EST dengan GDRP

H_1 : Ada hubungan antara HH_W dan EST dengan GDRP

Hasil pengolahan data uji hipotesis satu arah bahwa pengaruh kedua variabel prediktor yaitu HH_W dan EST adalah positif terhadap variabel respon GDRP, diberikan pada Tabel 6.2.3.3 berikut.

Tabel 6.2.3.3 Hasil pengolahan data *Income of all Districts and Cities in Indonesia*

LINEAR REGRESSION				
Number of observation : 501			F (2, 498) : 106.59	
		Prob > F	: 0.0000	
		R-squared	: 0.6561	
		Root MSE	: 1.7e+07	
GDRP	Coefficients	Robust Standard Error	t-test	P value > t
HH_W	176597.3	58217.43	3.03	0.003
EST	1347.857	100.6451	13.39	0.000
Constant	-9429857	2695074	-3.50	0.001

Source: processing data, 2020.

Dari tabel di atas terlihat bahwa nilai-p untuk kedua variabel prediktor yaitu HH_W dan EST masing-masing adalah 0,003 dan 0,000. Keduanya lebih kecil dari kriteria $\alpha = 0,05$. Maka kesimpulannya adalah kita menolak H_0, yang bermakna bahwa ada hubungan antara jumlah rumah tangga yang dapat mengakses air bersih di kabupaten dan kota (HH_W) dan jumlah guru SD di kabupaten dan di kota (EST) dengan Produk Domestik Regional Bruto (GDRP). Hasil pengujian ini membuktikan bahwa keberadaan air bersih dan guru SD berpengaruh terhadap Produk Domestik Bruto Daerah secara signifikan.

Nilai kedua parameter adalah positif, hal ini menunjukkan hasil penelitian ini sesuai dengan hipotesis dan landasan teori. Besarnya koefisien parameter mencerminkan besarnya pengaruh variabel bebas terhadap variabel terikat. Nilai parameter air bersih dengan koefisien 176,597 menunjukkan bahwa setiap kenaikan 1% akses rumah tangga terhadap air menghasilkan peningkatan sekitar Rp.176.597 juta di setiap kabupaten dan kota di seluruh Indonesia. Demikian juga untuk parameter guru SD sebesar 1.348 menunjukkan bahwa kenaikan satu orang guru SD menghasilkan kenaikan sekitar 1.348 juta rupiah di setiap kabupaten dan kota di Indonesia. Jadi

hasil penelitian ini juga menyarankan peningkatan akses air bersih dan jumlah dan kualitas guru sekolah dasar secara intensif untuk mencapai program pembangunan bangsa**.**

6.3. Uji Kolmogorov-Smirnov untuk Pengujian Kenormalan Data

Bila tersedia uji-uji statistik yang dapat digunakan dalam suatu rancangan penelitian, maka diperlukan untuk menggunakan dasar pemikiran tertentu untuk menentukan pilihan yang mana di antara uji-uji statistik yang akan dipakai. Bagian dari kekuatan suatu analisis statistika adalah suatu fungsi uji statistik yang dipakai dalam analisis tersebut. Suatu uji yang baik jika mempunyai peluang yang kecil untuk menolak H_0 apabila H_0 benar; yaitu taraf signifikansi uji yang dinyatakan dengan α. Kriteria uji yang lain adalah mempunyai peluang yang besar untuk menolak H_0 pada saat H_0 salah; yaitu γ yang dikenal sebagai fungsi kuasa uji (halaman).

Semua uji yang telah dibahas di atas merupakan uji statistik yang dikenal luas dan dikategorikan sebagai uji statistik parametrik, yaitu suatu uji yang dilandasi oleh asumsi bahwa distribusi data atau pengamatan menyebar menurut Distribusi Normal, dengan parameter nilai rata-rata μ dan standar deviasi σ tertentu. Uji statistik parametrik dikenal sebagai uji yang kuasa (*powerfull*) dalam pengertian memberikan nilai peluang γ yang besar. Tetapi apabila asumsi distribusi normal tidak dipenuhi, maka uji parametrik menjadi tidak sahih, dalam pengertian dapat memberikan kesimpulan yang sesat. Dari sini dikenal uji statistik nonparametrik yang merupakan uji statistik yang tidak dilandasi asumsi tersebut. Pada keadaan lain, data yang dihadapi hanya diukur pada skala nominal atau ordinal, sehingga tidak mungkin dianalisis dengan menggunakan uji statistik parametrik Dalam situasi demikian, hanya uji statistik nonparametrik yang dapat digunakan.

Dalam bagian ini kita akan membahas salah satu uji statistik nonparametrik yang banyak digunakan dalam penelitian yaitu uji kenormalan distribusi suatu kumpulan data. Uji ini dikenal sebagai Uji Kebaikan Suai (*Goodness of Fit Test*) yaitu suatu uji yang digunakan menguji hipotesis apakah suatu sampel acak yang ditarik dari suatu populasi menyebar sesuai dengan distribusi peluang teoritis tertentu. Dalam penggunaan uji statistik parametrik, distribusi yang dimaksudkan adalah Distribusi Normal.

Salah satu uji kenormalan distribusi yang umum digunakan adalah Uji Satu Sampel Kolmogorov-Smirnov. Uji ini akan menetapkan apakah skor-skor dalam sampel acak dapat secara masuk akal dianggap berasal dari suatu populasi dengan teoritis tertentu. Hipotesis nol H_0

adalah sampel acak berasal dari suatu populasi dengan teoritis tertentu dengan hipótesis alternatif H_1 adalah sampel acak tidak berasal dari suatu populasi dengan teoritis tertentu tersebut. Hipotesis ini dituliskan sebagai berikut.

H_0 : Distribusi sampel acak berasal dari suatu populasi dengan teoritis tertentu

H_1 : Distribusi sampel acak tidak berasal dari suatu populasi dengan teoritis di bawah H_0

Uji ini mencakup perhitungan distribusi frekuensi kumulatif data pengamatan yang akan terjadi di bawah distribusi teoritisnya (di bawah H_0) yaitu $F_0(x)$ dan membandingkannya dengan distribusi frekuensi kumulatif data pengamatan yaitu $F_N(x)$, di mana N adalah banyaknya sampel. Uji akan menetapkan suatu titik data pengamatan di mana kedua distribusi frekuensi kumulatif tersebut memiliki perbedaan yang terbesar. Jika distribusi peluang teoritis bersesuaian dengan distribusi sampel acak, maka tentu saja perbedaan ini sangat kecil; dan jika semakin besar maka dapat dikatakan distribusi peluang teoritis tidak bersesuaian dengan distribusi sampel acak. Jika dilihat dari nilai-p, apabila nilai-p > α maka hipotesis nol diterima yaitu sampel acak berasal dari suatu populasi dengan teoritis tertentu tersebut.

Uji Kolomogorov-Smirnov memusatkan perhatian pada penyimpangan (deviasi) terbesar, dan harga $F_0(x) - F_N(x)$ terbesar yang dinamakan statistik deviasi maksimum (D), yaitu

$$D = \text{maksimum} \mid F_0(x) - F_N(x) \mid$$

Distribusi sampel D di bawah H_0 diketahui, dan tersedia Tabel Nilai kritis Uji Kolmogorov-Smirnov (uji dwi-arah) (Tabel 6.3) yang memberikan nilai-nilai kritis tertentu dari distribusi sampling D. Perlu diketahui, bahwa signifikansi suatu nilai D tertentu bergantung pada N.

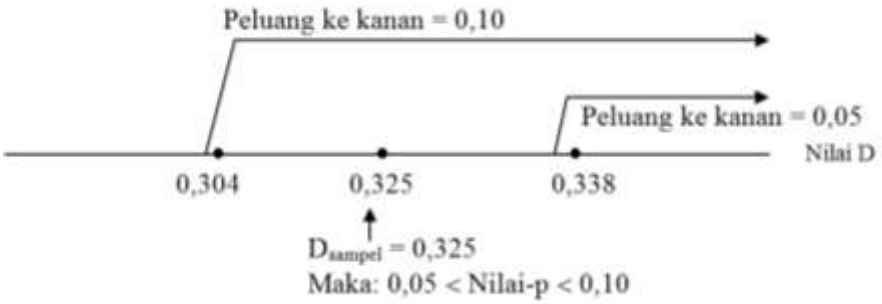
Contoh penggunaan Tabel 6.3.

Misalkan $D_{sampel} = 0{,}325$ dengan N = 15.

Dari tabel, $D \geq 0{,}325$ terjadi antara $\alpha = 0{,}05$ dan $\alpha = 0{,}10$ (uji dua-arah)

Maka nilai-p terletak antara 0,05 dan 0,10; atau lebih besar dari kriteria $\alpha = 0{,}05$.

Ilustrasi penentuan nilai-p dengan Tabel 6.3 diperlihatkan pada Gambar 6.3.1 berikut.



Gambar 6.3.1 Ilustrasi penentuan nilai-p Uji Kolmogorov-Smirnov

Jika N > 35 , nilai kritis D ditentukan secara khusus. Misal N = 43, dan $\alpha = 0{,}05$, maka untuk setiap $D \geq \frac{1{,}35}{\sqrt{43}} = 0{,}206$ akan signifikan pada tingkat $\alpha = 0{,}05$ (uji dua-arah)

Tabel 6.3 Nilai kritis Uji Kolmogorov-Smirnov

n	α = 0,20	α = 0,10	α = 0,05	α = 0,02	α = 0,01
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
6	0,410	0,468	0,519	0,577	0,617
7	0,381	0,436	0,483	0,538	0,576
8	0,359	0,410	0,454	0,507	0,542
9	0,339	0,387	0,430	0,480	0,513
10	0,323	0,369	0,409	0,457	0,486
11	0,308	0,352	0,391	0,437	0,468
12	0,296	0,338	0,375	0,419	0,449
13	0,285	0,325	0,361	0,404	0,432
14	0,275	0,314	0,349	0,390	0,418
15	0,266	0,304	0,338	0,377	0,404
16	0,258	0,295	0,327	0,366	0,392
17	0,250	0,286	0,318	0,355	0,381
18	0,244	0,279	0,309	0,346	0,371
19	0,237	0,271	0,301	0,337	0,361
20	0,232	0,265	0,294	0,329	0,352
21	0,226	0,259	0,287	0,321	0,344
22	0,221	0,253	0,281	0,314	0,337
23	0,216	0,247	0,275	0,307	0,330
24	0,212	0,242	0,269	0,301	0,323
25	0,208	0,238	0,264	0,295	0,317
26	0,204	0,233	0,259	0,290	0,311
27	0,200	0,229	0,254	0,284	0,305
28	0,197	0,225	0,250	0,279	0,300
29	0,193	0,221	0,246	0,275	0,295
30	0,190	0,218	0,242	0,270	0,290
35	0,177	0,202	0,224	0,251	0,269
40	0,165	0,189	0,210	0,235	0,252
45	0,156	0,179	0,198	0,222	0,238
50	0,148	0,170	0,188	0,211	0,226
55	0,142	0,162	0,180	0,201	0,216
60	0,136	0,155	0,172	0,193	0,207
65	0,131	0,149	0,166	0,185	0,199
70	0,126	0,144	0,160	0,179	0,192
75	0,122	0,139	0,154	0,173	0,185
80	0,118	0,135	0,150	0,167	0,179
85	0,114	0,131	0,145	0,162	0,174
90	0,111	0,127	0,141	0,158	0,169
95	0,108	0,124	0,137	0,154	0,165
100	0,106	0,121	0,134	0,150	0,161
Pendekatan					
n	1,07/√n	1,22/√n	1,35/√n	1,52/√n	1,63/√n
200	0,076	0,086	0,096	0,107	0,115

Sumber Tabel: https://luk.staff.ugm.ac.id/stat/ks/Kolmogorov-SmirnovTable.pdf

6.4. Artikel Ilmiah Menggunakan Uji Kolmogorov-Smirnov

6.4.1. Artikel Publikasi Jurnal Teknik Mesin – ITI Vol.6 No.1, Februari 2022

Samosir, P. dkk (2022) telah melakukan penelitian dengan judul Analisis Statistika Hubungan antara Lebar Pemotongan (*Kerf*) dengan Variabel Proses Permesinan *Wire-Electrical Discharge Machining* (WEDM). Percobaan menggunakan rancangan faktorial 3 faktor dengan replikasi sebanyak tiga kali dan analisis Statistika menggunakan model regresi dengan variabel prediktor adalah variabel-variabel proses WEDM yaitu variabel *arc on time* (AoT), variabel *servo voltage* (SV) dan variabel *wire feed* (WF), sedangkan variabel respon adalah lebar pemotongan (*kerf*). Penyusunan model regresi dilakukan dengan menggunakan pendekatan analisis variansi untuk rancangan faktorial dengan tiga faktor dengan semua variabel prediktor yaitu variabel proses WEDM merupakan variabel kualitatif. Salah satu asumsi yang mendasari penggunaan analisis ini bahwa distribusi variabel respon adalah normal. Oleh karenanya dalam contoh ini kita akan membahas teknik pemeriksaan asumsi kenormalan variabel respon yaitu lebar pemotongan (*kerf*). Data diperoleh dari hasil penelitian yang dilakukan Rupajati, P., dkk. (2018) dan diberikan pada Tabel 6.4.1 berikut ini.

Tabel 6.4.1. Data Lebar Pemotongan (*Kerf*)

No.	AoT (A)	SV (volt)	WF (mm/min)	Lebar Pemotongan (μm)
1	2	37	3	359
2	2	37	7	363
3	2	37	12	342
4	2	55	3	365
5	2	55	7	361
6	2	55	12	345
7	2	70	3	356
8	2	70	7	362
9	2	70	12	368
10	4	37	3	361
11	4	37	7	373
12	4	37	12	355
13	4	55	3	368
14	4	55	7	365
15	4	55	12	347
16	4	70	3	355
17	4	70	7	366
18	4	70	12	369
19	5	37	3	360
20	5	37	7	380
21	5	37	12	357
22	5	55	3	364
23	5	55	7	370

24	5	55	12	354
25	5	70	3	359
26	5	70	7	370
27	5	70	12	366

Hipotesis dalam Uji Kolmogorov-Smirnov ini dituliskan sebagai berikut.

H_0 : Distribusi data lebar pemotongan (*kerf*) berasal dari suatu populasi dengan Distribusi Normal

H_1 : Distribusi data lebar pemotongan (*kerf*) tidak berasal dari suatu populasi dengan distribusi Normal

Pengolahan data Uji Kolmogorov-Smirnov disajikan dalam Tabel 6.4.3 berikut dengan keterangan sebagai berikut:

1. Kolom 1 adalah data mentah Lebar Pemotongan (*Kerf*) dengan N = 27. Untuk meyusun distribusi frekuensi kumulatif, data mentah tersebut harus diurutkan terlebih dahulu.
2. Kolom 2 adalah data Lebar Pemotongan (*Kerf*) yang telah diurutkan.Terlihat ada beberapa data yang berulang, maka data disusun ke dalam 20 kelas (k = 20) sesuai dengan jumlah data yang berulang untuk penyusunan Frekuensi Kumulatif.
3. Kolom 3 adalah data untuk penyusunan Frekuensi Kumulatif S_N dengan banyak kelas k = 20.
4. Kolom 4 adalah Frekuensi Kumulatif S_N dengan banyak kelas k = 20.
5. Kolom 5 adalah data F_N pada kolom 4 dengan banyak kelas k = 20.
6. Kolom 6 adalah data F_N pada kolom 5 yang disusun kembali untuk data pada kolom 2 yaitu data dengan N = 27.
7. Kolom 7 adalah data nilai z, yang diperoleh dengan langkah berikut.
   a. Dengan menggunakan Program Minitab ditentukan rata-rata dan standar deviasi data mentah pada kolom 1 (Tabel 6.4.2).

Tabel 6.4.2. Statistik Lebar Pemotongan (Kerf)_1

Variabel	Rata-rata	Standar deviasi
Lebar Pemotongan (Kerf)_1	361,48	8,55

   b. Maka data pada kolom 7 diperoleh dengan rumus transformasi Normal Baku:

$$z = \frac{\text{Data kolom 2} - (\text{Rata} - \text{rata})}{\text{Standard deviasi}} = \frac{\text{Data kolom 2} - 361{,}48}{8{,}55}$$

8. Kolom 8 adalah nilai peluang kumulatif di bawah Distribusi Normal yang dihitung dengan menggunakan menu *Cumulative Probability* Program Minitab, yaitu:

   $F_0(z) = P(Z < z)$

9. Kolom 9 adalah nilai $| F_0(x) - F_N(x) |$ dari data Kolom 6 dan Kolom 8.

Tabel 6.4.3. Pengolahan Data Uji Kolmogorov-Smirnov

Data Mentah Lebar Pemotongan (Kerf)	Data Terurut Lebar Pemotongan (Kerf)	Data Untuk Penyusunan Frekuensi Kumulatif S_N	S_N (20 Kelas)	Distribusi Frekuensi Kumulatif F_N (20 Kelas)	F_N (27 Kelas)	Nilai z Data Kolom 2 (Dist.Normal)	F_0 (27 Kelas)	$\| F_0 - F_N \|$ (27 Kelas)
1	2	3	4	5	6	7	8	9
359	342	342	1	0.037037	0.037037	-2.27836	0.011352	0.025685
363	345	345	2	0.074074	0.074074	-1.92749	0.02696	0.047114
342	347	357	3	0.111111	0.111111	-1.69357	0.045174	0.065937
365	354	354	4	0.148148	0.148148	-0.87485	0.190827	0.042679
361	355	355	6	0.222222	0.222222	-0.75789	0.224257	0.002035
345	355	356	7	0.259259	0.222222	-0.75789	0.224257	0.002035
356	356	357	8	0.296296	0.259259	-0.64094	0.260782	0.001523
362	357	359	10	0.37037	0.296296	-0.52398	0.300147	0.003851
368	359	360	11	0.407407	0.37037	-0.29006	0.385886	0.015515
361	359	361	13	0.481481	0.37037	-0.29006	0.385886	0.015515
373	360	362	14	0.518519	0.407407	-0.1731	0.431287	0.023879
355	361	363	15	0.555556	0.481481	-0.05614	0.477615	0.003866
368	361	364	16	0.592593	0.481481	-0.05614	0.477615	0.003866
365	362	365	18	0.666667	0.518519	0.060819	0.524248	0.00573
347	363	366	20	0.740741	0.555556	0.177778	0.570551	0.014996
355	364	368	22	0.814815	0.592593	0.294737	0.615903	0.02331
366	365	369	23	0.851852	0.666667	0.411696	0.659719	0.006948
369	365	370	25	0.925926	0.666667	0.411696	0.659719	0.006948
360	366	373	26	0.962963	0.740741	0.528655	0.701478	0.039263
380	366	380	27	1	0.740741	0.528655	0.701478	0.039263
357	368				0.814815	0.762573	0.777141	0.037674
364	368				0.814815	0.762573	0.777141	0.037674
370	369				0.851852	0.879532	0.810444	0.041408
354	370				0.925926	0.996491	0.840494	0.085432
359	370				0.925926	0.996491	0.840494	0.085432
370	373				0.962963	1.347368	0.911069	0.051894
366	380				1	2.166082	0.984848	0.015152

Maka statistik uji Kolmogorov-Smirnov adalah:

$D = \text{maksimum} \, | F_0(x) - F_N(x) | = 0{,}085432 \approx 0{,}09.$

Nilai-p ditentukan dengan menggunakan Tabel Uji Kolmogorov-Smirnov. Berdasarkan Tabel 6.3 untuk N = 27, nilai kritis terkecil untuk menolak H_0 adalah 0,2 dengan taraf signifikansi 0,2. Semakin kecil nilai taraf signifikansi, semakin besar nilai kritis D. Maka untuk D_{sampel} = 0,09 diperkirakan nilai-p lebih besar dari taraf signifikansi 0,2. Dengan demikian kita menerima H_0

bahwa distribusi data lebar pemotongan (*kerf*) berasal dari suatu populasi dengan Distribusi Normal.

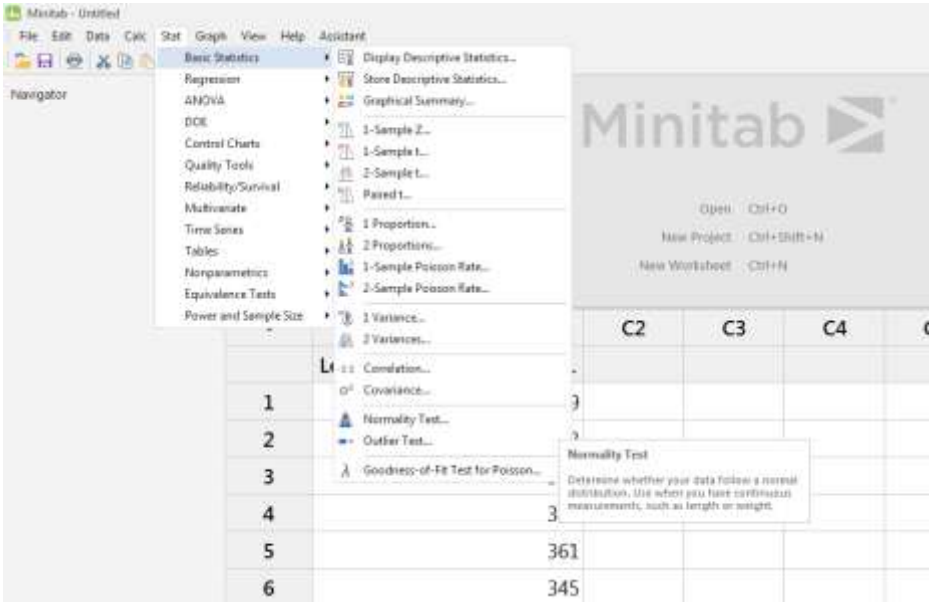
Berikut ini langkah uji Kolmogorov-Smirnov dengan menggunakan Program Minitab.

Langkah 1. Input data pada kolom C1. Diberi nama Lebar Pemotongan (Kerf)_1, sebab data ini akan melewati beberapa proses pengolahan data seperti pada Tabel di atas dan harus dengan nama berbeda (Gambar 6.4.1.1).

	C1	C2	C3	C4	C5
	Lebar Pemotongan (Kerf)_1				
1	359				
2	363				
3	342				
4	365				
5	361				
6	345				
7	356				
8	362				
9	368				
10	361				
11	373				

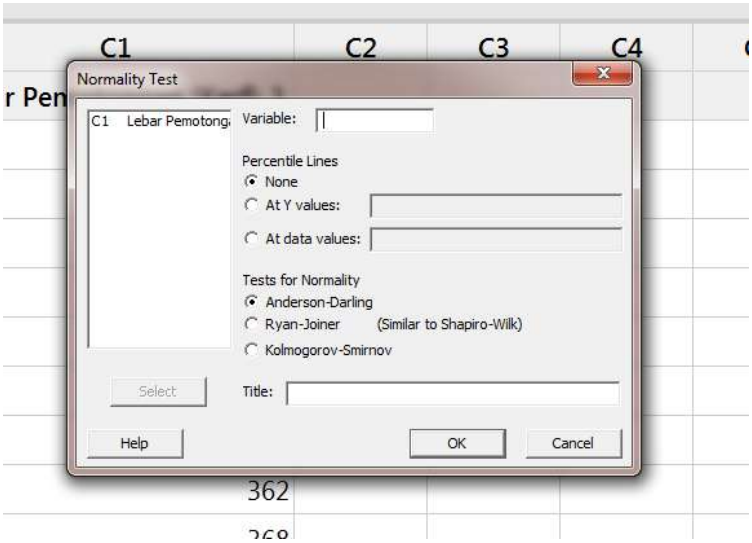
Gambar 6.4.1.1 Input data Lebar Pemotongan (Kerf)_1

Langkah 2. Dari ikon *Stat*, pilih *Basic Statistics*, kemudian *Normality Test* (Gambar 6.4.1.2).



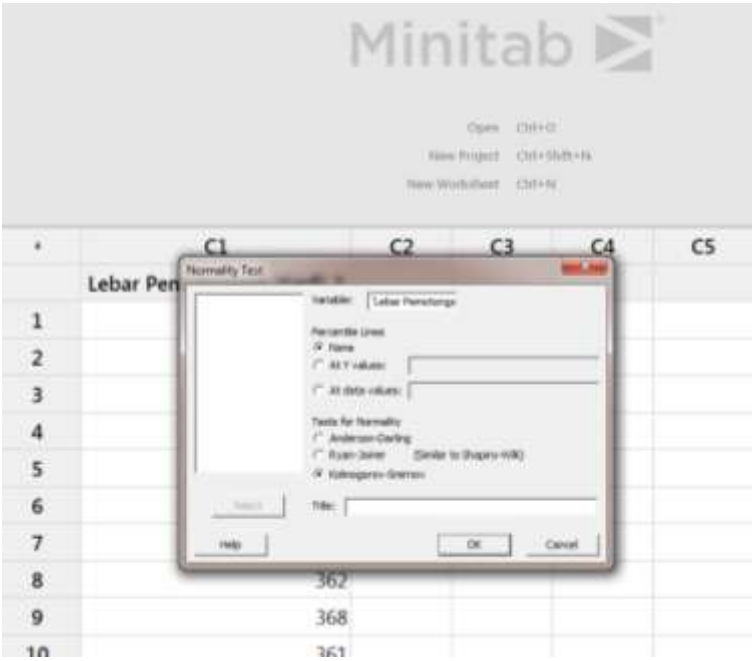
Gambar 6.4.1.2 Menu *Normality Test*

Langkah 3. Akan muncul kotak dialog *Normality Test*. Terlihat data C1 Lebar Pemotongan (*Kerf*) di sebelah kiri. Kemudian ada 3 pilihan uji kenormalan (*Tests for Normality*) dan *defaultnya* adalah Anderson-Darling (Gambar 6.4.1.3).



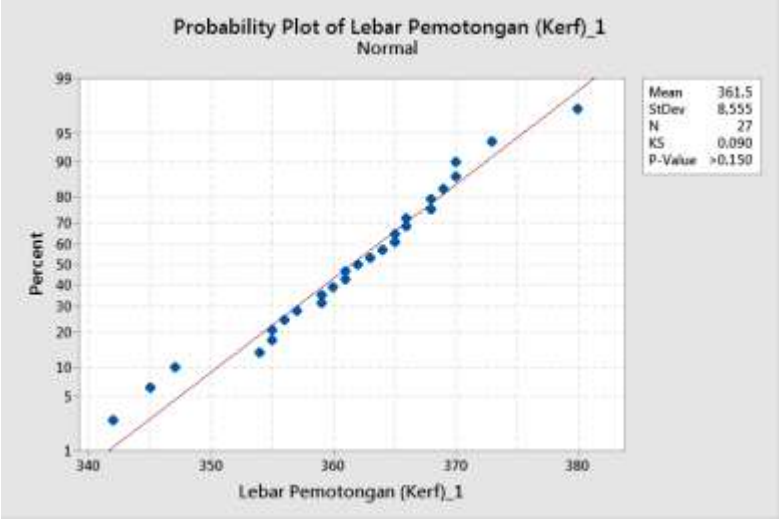
Gambar 6.4.1.3 Kotak dialog *Normality Test*

Langkah 4. Pindahkan data C1 Lebar Pemotongan (*Kerf*) ke kotak *Variable* dan pilih *Tests for Normality*: Kolmogorov-Smirnov (Gambar 6.4.1.4).



Gambar 6.4.1.4 Opsi Uji Kolmogorov-Smirnov

Langkah 5. Klik Ok dan hasilnya diberikan pada Gambar 6.4.1.5 berikut.



Gambar 6.4.1.5 Hasil pengujian hipotesis data lebar pemotongan (*kerf*)

Dari Gambar 6.4.1.5 di atas terlihat nilai statistik uji Kolmogorov-Smirnov yaitu KS = 0,090 dengan P-value > 0,150. Dengan demikian kita menerima hipotesis nol H_0 bahwa distribusi data lebar pemotongan (*kerf*) pada sampel ini berasal dari suatu populasi dengan Distribusi Normal.

6.4.2. Artikel Publikasi Jurnal IPTEK Volume 1 Nomor 1, April 2015

Uji Kolmogorov-Smirnov untuk memeriksa kenormalan distribusi dalam penggunaan Analisis Regresi juga dilakukan dalam penelitian Perak Samosir dan Ratnawati (2015) tentang Pengaruh Derajat Infeksi Cacing terhadap Tingkat Kecerdasan Anak (Studi Kasus terhadap Siswa SD Negeri 067775 Kotamadya Medan). Artikel hasil penelitian ini telah dipublikasikan Jurnal IPTEK Volume 1 Nomor 1, April 2015: 7 – 12.

Infeksi cacing usus yang ditularkan melalui tanah (*soil transmitted helminthiasis*/STH) dapat mempengaruhi kualitas sumber daya manusia melalui gangguan penyerapan gizi yang dapat menyebabkan penurunan tingkat intelegensia anak. Maka di dalam penelitian ini dilakukan Analisis Statistika untuk melihat besar pengaruh derajat infeksi, kadar hemoglobin dan status gizi terhadap tingkat kecerdasan anak pada siswa SD Negeri 067775 Kotamadya Medan. Analisis dilakukan dengan menggunakan Analisis Regresi Linear Berganda. Variabel respon adalah tingkat kecerdasan siswa dan variabel prediktor adalah derajat infeksi, kadar hemoglobin dan status gizi.

Data pada penelitian ini merupakan data sekunder. Data diperoleh dari hasil penelitian yang dilakukan oleh Agustina Samosir di SD Negeri di Medan yakni SD Negeri 067775 Kotamadya Medan, dari bulan Juli 2012 sampai dengan Agustus 2012. Sampel penelitian diambil dari populasi anak–anak SD Negeri 067775. Data variabel respon (tingkat intelegensi) dari 62 siswa diberikan dalam Tabel 6.4.2 berikut.

Tabel 6.4.2 Data tingkat intelegensi siswa

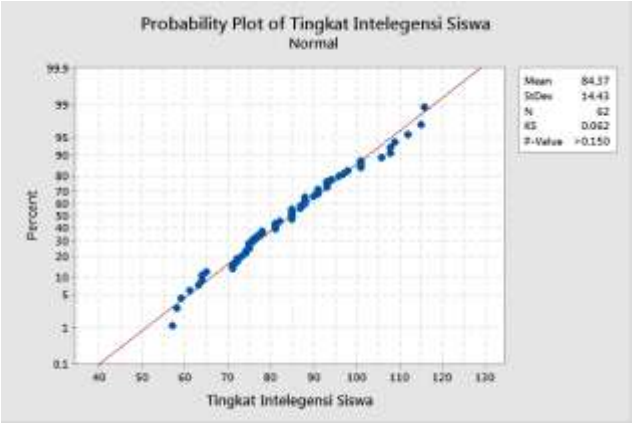
91	85	65	108	85	101
98	72	88	93	81	109
58	57	77	101	74	64
108	75	97	81	88	85
116	78	88	115	87	77
73	64	112	71	93	82
72	59	106	81	90	81
71	101	96	76	93	76
85	75	63	88	78	85
91	61	94	85	87	74
75	91				

Dengan Uji Kolmogorov-Smirnov akan diuji kenormalan data tingkat intelegensi siswa. Rumusan hipotesisnya diberikan sebagi berikut.

H_0 : Data tingkat intelegensi siswa berdistribusi Normal

H_1 : Data tingkat intelegensi siswa bukan berdistribusi Normal

Dikarenakan ukuran sampel cukup besar (n = 62), maka prosedur pengujian yang disajikan di sini hanya hasil uji dengan menggunakan Program Minitab, dan hasilnya diberikan dalam Gambar 6.4.2 berikut.



Gambar 6.4.2 Hasil uji kenormalan data tingkat intelegensi siswa

Dari Gambar 6.4.2 di atas terlihat nilai statistik uji Kolmogorov-Smirnov yaitu KS = 0,052 dengan P-value > 0,150. Dengan demikian kita menerima hipotesis nol H_0 bahwa distribusi data data tingkat intelegensi siswa pada sampel ini berasal dari suatu populasi dengan Distribusi Normal.

DAFTAR PUSTAKA

Andreas, K., Suastiyanti, D. & Rupajati, P. Peningkatan Daya Listrik Pada Generator Putaran Rendah Melalui Peningkatan Sifat Magnetik Magnet Permanen BaFe12O19. Vol 4, No 1 2020: *Jurnal Teknik Mesin ITI*. 2020.

Budiono, S. 2009. Investasi dan Pertumbuhan Ekonomi di Indonesia (Penekanan Pada Investasi Pendidikan). *Jurnal Studi Ekonomi*, IV, 2009. 123 – 129.

Budiono, S., Purba, J. T., Rajagukguk, W., Samosir, P & Adirinekso, G. P. 2020. Measurement Availability of Clean Water and Elementary Teachers towards Income of all Districts and Cities in Indonesia. *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management*, Detroit, Michigan, USA, August 10 - 14, 2020. 1552 – 1559.

Cann, O. 2016. *These Are The Top 10 Emerging Technologies of 2016*. Annual Meeting of the New hampions. https://www.weforum.org/agenda/2016/06/top-10-emerging-technologies-2016/

Hogg, R. V., McKean, J. W. & Craig, A. T. 2013. *Introduction to Mathematical Satistics*. Pearson Education, Inc.

https://www.bps.go.id/menu/1/sejarah.html
https://luk.staff.ugm.ac.id/stat/ks/Kolmogorov-SmirnovTable.pdf

https://accounting.binus.ac.id/2021/08/13/memahami-uji-hausman-test-dalam-software-pengolahan-data-eviews/

Mendenhall, W. & Sincich, T. 2012. *A Second Course in Statistics: Regression Analysis*. Seventh Edition. Pearson Education, Inc.

Purba, J. T. & & Tan, J. D. 2016. Strategic Service Management in Hospitality Industry: Case Study of Grand Tasty Restaurants. *Conference on Management and Behavioral Studies* Universitas Tarumanagara, Jakarta, 27 Oktober 2016.

Purba, J. T. & Budiono, S. 2019. Availability of Electricity, Clean Water and Sanitation towards Economic Growth in Indonesia with 500 Regencies and Cities. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Riyadh, Saudi Arabia. 657–662.

Purba, J. T., Budiono, S, Rajagukguk, W., Samosir, P. & Adirinekso,G. P. 2020. E-Business Services Strategy with Financial Technology: Evidence from Indonesia. *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management*, Detroit, Michigan, USA, August 10 - 14, 2020. 1543 – 1551.

Rajagukguk, W., 2016. The impact of joining wto on indonesia's economy: Econometric Modelling Approach. *Actual Problems of Economics*. 176. 59 – 67.

Rajagukguk, W. & Purba, JT. (2012). Perekonomian Indonesia dan Strategi Pembangunan Jender. *Jurnal Pengembangan Bisnis & Manajemen* 20 (Vol. XII), 24 – 29

Rajagukguk, W., Samosir, P., Purba, J. T., Budiono, S & Adirinekso, G. P. 2020. Dark Side of Economic Growth: A Case Study of the Relationship between Economic Growth and Suicide Mortality. *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management*, Detroit, Michigan, USA, *August* 10 - 14, 2020. 1474 – 1482.

Samosir, P & Ratnawati. 2015. Pengaruh Derajat Infeksi Cacing terhadap Tingkat Kecerdasan Anak (Studi Kasus terhadap Siswa SD Negeri 067775 Kotamadya Medan). *Jurnal IPTEK* Volume 1 Nomor 1, April 2015: 7 – 12.

Samosir, P. & Rajagukguk, W. 2020. Infrastructure Development and Economic Growth in Indonesia: A Province Panel Data Analysis. *International Journal of Innovation, Creativity and Change*. www.ijicc.net Volume 14, Issue 10, 2020. 309 – 318.

Samosir, P & Simamora, D. B. S. 2020. Analisis Statistika Pengaruh Proses Sintering Terhadap Sifat Magnetik Material Multiferoik Berbasis Ferit. *Seminar Technopex ITI* – 2020.

Samosir, P., Suastiyanti, D., Rajagukguk, W., Purba, J. T & Budiono, S. 2021. Statistical Analysis of the Relationship between the Sintering Process and Magnetic Energy on the Improvement of the Magnetic Properties of the Permanent Magnet $BaFe_{12}O_{19}$. *Proceedings of the International Conference on Industrial Engineering and Operations Management* Monterrey, Mexico, November 3-5, 2021. 1676 – 1685.

Samosir, P., Rupajati, P. & Fahmi, I. N. 2022. Analisis Statistika Hubungan antara Lebar Pemotongan (*Kerf*) dengan Variabel Proses Permesinan *Wire-Electrical Discharge Machining* (WEDM). *Jurnal Teknik Mesin – ITI* Vol.6 No.1, Februari 2022. 33 – 42.

Sarwoko. *Statistik Inferensi untuk Ekonomi dan Bisnis*. 2007. Penerbit Andi Yogyakarta.

Sembiring, R. K. 1995. *Analisis Regresi*. Edisi Kedua . Penerbit ITB Bandung.

Suastiyanti, D, Maulida Y. N., Wijaya, M. 2019. Improving Magnetic Properties of BiFeO3-BaFe12O19 Solid Solution by Different Sintering Time and Temperatures of Sol-Gel Method, vol. 7, Issue 5, 2019, *Asian Journal of Applied Sciences*. ISSN: 2321 − 0893. Volume 07 – Issue 05, October 2019.

Tan, J. D., Purba, J. T. & Wijaya, A. E. 2019. Financial Technology as an Innovation Strategy for Digital Payment Services in the Millenial Generation. *Proceedings of the 1st Aceh Global Conference* (AGC 2018). 364-373.

Wackerly, D., Mendenhall III, W. Scheaffer, R. 2008. *Mathematical Statistics with Applications*. Ed 7th. Thomson Learning, Inc.

Walpole, R. E. & Myers, R. H. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*. Edisi ke-4, Penerbit ITB Bandung.

Wooldridge, J. M. 2002. Econometric Analysis of Cross Section and Panel Data. In *Booksgooglecom*. https://doi.org/10.1515/humr.2003.021

I N D E K S

Analisis Regresi
- Peubah respon, 158
- Peubah prediktor, 158
- Acak, 159
- Komponen deterministik, 159
- Komponen galat acak, 159

Asumsi Analisis Regresi
- $\varepsilon_i \sim N(0, \sigma^2)$, 164
- $cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j$, 164
- y_i dan y_j tidak berkorelasi, $i \neq j$, 164
- $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$, 164
- Penaksir tak bias untuk σ^2, 165
- Inferensi parameter β_0 dan β_1, 165

Budiono, S. dkk
- *Business cycle theory*, 180
- Kesejahteraan ekonomi lokal, 180
- Ketersediaan air bersih, 180
- Guru SD, 180
- Data *cross section* 501 kabupaten dan kota di Indonesia, 180
- Purba, J. T. & Budiono, S., 180
- Budiono, S., 180
- Jumlah rumah tangga, 180
- Jumlah guru SD, 180
- Produk Domestik Regional Bruto, 180
- Hasil pengolahan data *Income of all Districts and Cities in Indonesia*, 181

Dasar-dasar teori, 3

Data
- Data primer, 1
- Data sekunder, 1

Distribusi Binomial
- Tabel Peluang Binomial, 26
- Tabel Distribusi Kumulatif Binomial, 26
- Penggunaan Program Minitab, 26
- Menu Distribusi Binomial, 27
- Input data *Probability*, 28
- *Output Probability* Binomial, 30
- *Output Cumulative probability* Binomial, 31
- Grafik Distribusi Binomial, 32

Distribusi F, 51

Distribusi Hipergeometrik, 33

Distribusi Khi-kuadrat
- Kotak dialog *Calc*, 45
- *Expression* untuk $1 - \alpha$, 46
- *Output Expression* $1 - \alpha$, 46

Distribusi Normal, 39
- z-*score*, 42
- *Inverse cumulative probability*, 43

Distribusi peluang dan grafik, 12

Distribusi peluang teoritis, 24

Distribusi Poison, 37

Distribusi t, 49

Distribusi Seragam, 25

Distribusi t, 49

Ekspektasi, 22

Ilmu Statistika, 1

Kejadian saling bebas, 18

Kejadian saling meniadakan, 16

Kriteria penaksir
- Penaksir tak bias, 61
- Penaksir bias, 63
- Teorema *Weak Law of Large* Numbers, 63
- Teorema Chebyshev's *Inequality*, 64
- Penaksir konsisten, 65

Konsep pengujian hipotesis
- Galat jenis I, 91
- Galat jenis II, 91
- Daerah kritis, 92
- Peluang α, β dan γ, 92
- Uji ekaarah, 97
- Uji dwiarah, 99
- Langkah pengujian hipotesis, 101

Konsep ukuran ketepatan Model Regresi
- Simpangan amatan, 162
- Simpangan prediksi, 162
- JKG, JKR dan JKT, 162
- Koefisien determinasi, 163
- Derajat bebas, 163
- Tabel ANOVA, 163

Model Regresi Linear Berganda
- Persamaan, 165
- Data Model RLB, 166
- JKG Model RLB, 166
- Penentuan penaksir β_i , $i = 0, 1, 2, \ldots, k$, 167
- Inferensi parameter β_i, $i = 0, 1, 2, \ldots, k$, 167

Model Regresi Linear Sederhana
- Parameter Model Regresi, 160
- Dugaan/penaksir/estimasi, 160
- Pengamatan RLS, 161
- Jumlah Kuadrat Galat (JKG), 161
- Penentuan penaksir untuk β_0 dan β_1, 161

Parameter, 54
Peluang, 5
Peluang bersyarat, 13
Peluang suatu kejadian, 8
Penaksiran, 54
Penaksir kemungkinan maksimum
Fungsi kemungkinan, 56
Persamaan estimasi, 56
Kasus 1 parameter, 57
Kasus 2 parameter, 58
Penaksiran parameter
Karakteristik data, 66
Taksiran titik, 66
Taksiran selang, 66
Pengujian hipotesis rata-rata dan variansi
Rata-rata suatu populasi (n≥30), 103
Rata-rata suatu populasi (n<30), 114
Selisih rata-rata 2 populasi bebas ($n_1 \geq 30$ dan $n_2 \geq 30$), 120
Selisih rata-rata 2 populasi bebas ($n_1 < 30$ dan $n_2 < 30$), 122
Selisih rata-rata pengamatan berpasangan, 129
Variansi suatu populasi, 133
Rasio variansi dua populasi, 137
Pengolahan data, 1
Percobaan
Ruang sampel, 5
Kejadian, 5
Himpunan, 6
Perhitungan jumlah titik sampel
Aturan perkalian, 10
Kombinasi, 10
Kombinasi dan Aturan Perkalian, 11
Peubah acak, 20
Peubah acak diskrit, 21
Peubah acak kontinu, 22
Purba, dkk
E-Business Services Strategy, 177
Tan, J. D. dkk., 178
Aplikasi *Fintech*, 178
Gopay, 178
P.T. Gojek Indonesia, 178
Cann, 178
Preferensi pengguna, 178
Layanan *fintech*, 178
Manfaat *fintech*, 178
Keamanan *fintech*, 178
GoPay *Fintech*, 178
Regresi *Robust*, 178
Outlier, 179
Stata, 179
Hasil pengolahan data *Fintech*, 179
Rajagukguk, W. dkk
Dark Side of Economic Growth, 175
Suicide mortality rate, 175
Female suicide mortality rate, 175
Male suicide mortality rate, 175
Pertumbuhan Produk Domestik Bruto, 175
Regresi Linier Berganda Multivariat, 176
Hasil pengolahan data, 177
Rajagukguk, W & Purba, J.T., 177
Samosir, P. dkk
Samosir, P dkk., 168
Suastiyanti, D. dkk., 168
Andreas, K. dkk, 168
Samosir, P & Simamora, D. B. S., 168
Barium heksaferit, 168
Proses sintering, 168
Sifat magnetik, 168
Energi magnetik ($\widehat{EM}$), 168
Persamaan $\widehat{EM}$, 170
Hasil uji hipotesis $\widehat{EM}$, 170
Samosir, P & Rajagukguk W
Kesenjangan pertumbuhan ekonomi, 171
Rajagukguk, W, 171
Data panel, 171
Fixed Effect Model (FEM), 172
Data *cross section*, 172
Data *time series*, 172
Persamaan *Fixed Effect Model*, 173
Produk Domestik Bruto, 173
Akses Listrik Rumah Tangga, 173
Angka Melek Huruf, 173
Persentase Kelahiran, 173
Hasil uji Hausmann, 174
Persamaan pertumbuhan ekonomi, 175
Samosir, P. dkk
Proses WEDM, 185
Arc on time, 185
Servo voltage, 185
Wire feed, 185

Lebar pemotongan (*kerf*), 185
Rupajati, P. dkk, 185
Pengolahan data Uji KS, 187
Uji KS dengan Program Minitab, 188
Menu *Normality test*, 188
Hasil uji kenormalan data
lebar pemotongan (*kerf*), 190
Samosir, P & Ratnawati
Derajat infeksi cacing, 190
Tingkat kecerdasan anak, 190
Kadar hemoglobin, 190
Status gizi, 190
Data tingkat intelegensi siswa, 191
Hasil uji kenormalan data tingkat
intelegensi siswa, 191
Sampel acak, 55
Selang kepercayaan $(1 - \alpha)100\%$
Rata-rata suatu populasi ($n \geq 30$), 68
Rata-rata suatu populasi ($n < 30$), 70
Selisih rata-rata 2 populasi bebas
($n_1 \geq 30$ dan $n_2 \geq 30$), 72
Selisih rata-rata 2 populasi bebas
($n_1 < 30$ dan $n_2 < 30$), 75
Selisih rata-rata pengamatan berpasangan,
79
Proporsi suatu populasi, 81
Hampiran Normal terhadap Binomial, 82
Selisih proporsi 2 populasi, 84
Variansi suatu populasi, 86
Rasio variansi dua populasi, 88
Statistik, 54
Statistika Deskriptif, 1
Statistika Inferensi, 2
Uji hipotesis proporsi
Proporsi suatu populasi, 142
Menu 1 *Proportion*, 144
Selisih proporsi 2 populasi, 153
Menu *Two Proportions*, 155
Uji Kolmogorov-Smirnov
Taraf signifikansi uji, 182
Kuasa uji, 182
Uji Kebaikan Suai, 182
Deviasi maksimum, 183
Tabel Nilai kritis
Uji Kolmogorov-Smirnov (KS), 184
Teorema Limit Pusat, 48

BIORAFI PENULIS BUKU REFERENSI

Dra. Perak Samosir, M.Si. adalah Dosen Matematika I, Matematika II dan Statistik pada Program Studi Teknik Mesin Institut Teknologi Indonesia. Perak Samosir juga merupakan Dosen Ilmu Statistika (Teori Peluang, Statistika Matematika, Analisis Regresi dan Statistika Nonparametrik) pada Program Studi Matematika Fakultas Sains dan Teknologi (FAST) Universitas Pelita Harapan Indonesia. Gelar pertamanya di bidang Matematika dari Institut Teknologi Bandung dan gelar masternya di bidang Statistik dari Institut Pertanian Bogor Indonesia. Minat penelitiannya adalah pemodelan dengan menggunakan Analisis Statistika.

Dr. Wilson Rajagukguk, S.Si., M.Si., M.A. adalah Dosen Matematika Ekonomi, Ekonometrika dan Statistika di Fakultas Ekonomi. Master of Art bidang Theologia (M.A) diperoleh dari Sekolah Tinggi Theologia dan Magister Sains Ilmu Ekonomi bidang kependudukan dan ketenagakerjaan (M.Si) dan Doktor Ilmu Ekonomi bidang Ekonomi Regional (Dr.) dari Universitas Indonesia. Minat penelitiannya adalah di bidang pertumbuhan ekonomi demografi, ekonomi agama, statistika, ekonometrika, dan matematika.

Dra. Ratnawati, M. Si. adalah Dosen Kalkulus, Matematika Teknik I dan Matematika Teknik II pada Program Studi Teknik Elektro Institut Teknologi Indonesia. Gelar pertamanya di bidang Matematika dari Institut Teknologi Bandung dan gelar masternya di bidang Statistik dari Institut Pertanian Bogor Indonesia. Minat penelitiannya adalah pemodelan dengan menggunakan Analisis Statistika.

UKI PRESS
Pusat Penerbit dan Percetakan
Universitas Kristen Indonesia
Jl.Mayjen Sutoyo No. 02 Cawang
Jakarta Timur 13630

ISBN 978-623-6963-93-7