

# **APLIKASI MODEL LOGLINEAR DAN REGRESI LOGISTIK**

(Studi Kasus : Jumlah kasus Covid-19 pada tahun 2020)

## **TUGAS AKHIR**

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana  
Program Studi Statistika



Ika Meiza Maharani

16611053

**PROGRAM STUDI STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS ISLAM INDONESIA  
YOGYAKARTA  
2020**

## HALAMAN PERSETUJUAN PEMBIMBING

### TUGAS AKHIR

Judul : Aplikasi Model Loglinear dan Regresi Logistik  
(Jumlah kasus Covid-19 pada tahun 2020).

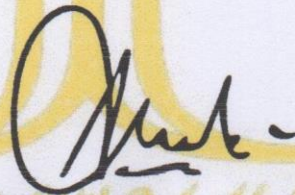
Nama Mahasiswa : Ika Meiza Maharani

NIM : 16611053

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK  
DIUJIKAN**

Yogyakarta, 10 Juli 2020

Pembimbing



(Jaka Nugraha, Dr., S.Si., M.Si.)

## HALAMAN PENGESAHAN

### TUGAS AKHIR

#### APLIKASI MODEL LOGLINEAR DAN REGRESI LOGISTIK

(Studi Kasus : Jumlah kasus Covid-19 pada tahun 2020)

Nama Mahasiswa : Ika Meiza Maharani

NIM : 16611053

TUGAS AKHIR INI TELAH DIUJIKAN  
PADA TANGGAL: 20 Juli 2020

Nama Penguji:

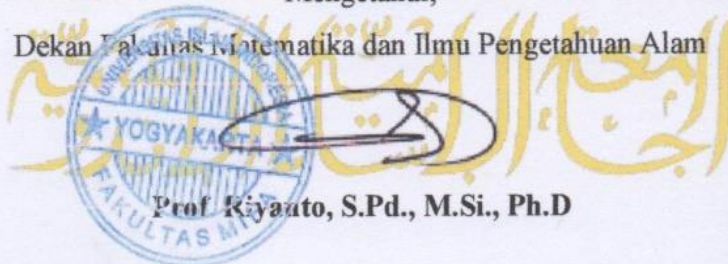
1. Edy Widodo, Dr., S.Si., M.Si.
2. Dr. Techn. Rohmatul Fajriyah, M.Si.
3. Jaka Nugraha, Dr., S.Si., M.Si.

Tanda Tangan



Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Prof. Kiyanto, S.Pd., M.Si., Ph.D

## KATA PENGANTAR



*Assalamu'alaikum Wr.Wb*

Puji Syukur peneliti panjatkan ke-hadirat Allah SWT, karena atas limpahan rahmat dan hidayahNyalah, sehingga peneliti dapat menyelesaikan Tugas Akhir yang berjudul “Aplikasi Model Loglinear dan Regresi Logistik” dengan menggunakan studi kasus Covid 19 pada tahun 2020. Shalawat serta salam tercurah kepada junjungan kita Nabi Muhammad SAW beserta keluarga dan para pengikut-pengikutnya.

Dalam Penyusunan Tugas Akhir ini, peneliti menyadari bahwa masih banyak kekurangan. Oleh karena itu, peneliti sangat mengharapkan kritik dan saran yang berkaitan dengan Tugas Akhir ini. Selama menyusun Tugas Akhir ini, peneliti telah banyak mendapat bimbingan dan bantuan dari berbagai pihak. Untuk itu pada kesempatan ini peneliti bermaksud menyampaikan ucapan terima kasih yang sebesar-besarnya kepada :

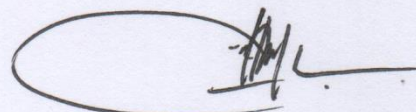
1. Bapak Prof.Riyanto, S.Pd.,M.Si.,Ph.D. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia, Yogyakarta.
2. Bapak Dr. Edy Widodo, S.Si.,M.Si., selaku Ketua Jurusan Statistika beserta seluruh jajarannya.
3. Bapak Jaka Nugraha, Dr., S.Si., M.Si. selaku dosen pembimbing yang telah memberi bimbingan peneliti selama penyusunan Tugas Akhir ini.
4. Bapak Poniran, Ibu Cucu Suhaeni, Adik tersayang Siwi Yulinda Amelia Putri dan Keluarga Besar yang selalu mendoakan yang terbaik untuk saya.
5. Mita Rahmatul Watania, Ade Arianti Putri dan Nur Isma Mardhotillah orang-orang yang paling penting dalam selesainya Tugas Akhir ini, sekaligus sebagai sahabat-sahabat yang kusayangi.
6. David Muhammad Musafie yang menjadi partner yang membantu memotivasi sekaligus menemani setiap hari.

7. Nurlia, Putri Atun, Mardhatillah dan Novia Rizki P yang menjadi sahabat terbaik selama di Yogyakarta. Terima kasih sudah selalu ada dalam senang maupun susah.
8. Orang ini ini aja, khususnya alm. Agus Hertanto, Bayu Galih P, Ristanti Nurmaulida dll yang telah menyemangati, dan selalu ada.
9. Personil sundae sebagai sahabat seperjuangan di Yogyakarta yang juga sebagai motivasi saya dalam menyelesaikan tugas akhir ini dengan cepat.
10. Teman seperbimbingan, khususnya untuk Anita G, Nada yang selalu memotivasi dan memberikan semangat sekaligus membantu dalam membuat laporan ini.
11. Sahabat-sahabat Statistika 2016 yang membantu dan menyemangati untuk menyelesaikan Tugas Akhir ini.
12. Serta semua pihak, baik secara langsung maupun tidak langsung telah membantu peneliti dalam menyelesaikan Tugas Akhir.

Semoga Allah SWT senantiasa melimpahkan rahmat dan hidayah-Nya kepada semua pihak yang telah membantu peneliti. Demikian Tugas Akhir ini, semoga bermanfaat bagi semua pihak. Amiin.

*Wassalamualaikum Wr.Wb*

Yogyakarta, 10 Juli 2020



Ika Meiza Maharani

## DAFTAR ISI

HALAMAN Sampul .....	i
HALAMAN PERSETUJUAN PEMBIMBING .....	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR .....	iv
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR .....	x
DAFTAR LAMPIRAN.....	xi
PERNYATAAN.....	xii
ABSTRAK .....	xiii
ABSTRACT.....	xiv
BAB I.....	1
PENDAHULUAN .....	1
1.1 Latar Belakang Masalah .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah .....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB II.....	5
TINJAUAN PUSTAKA .....	5
BAB III .....	8
LANDASAN TEORI.....	8
3.1 Coronavirus Disease (Covid-19) .....	8
3.2 Tabel Kontigensi Dua Dimensi .....	8
3.3 Regresi Logistik.....	9
3.3.1 Model Regresi Logistik .....	9
3.3.2 Penaksiran Parameter Model Logistik.....	10
3.3.3 Uji Signifikansi Parameter Logistik .....	11
3.3.4 Uji Kecocokan Model Logistik .....	12

3.3.5	Interprestasi Parameter Logistik .....	13
3.4	Model Loglinear .....	14
3.4.1	Model Loglinear dalam tabel kontigensi dua arah .....	14
3.4.2	Penaksiran Parameter Model Loglinear.....	17
3.4.3	Uji Kecocokan Model Loglinear .....	18
3.4.4	Penyeleksi Model Terbaik .....	19
3.4.5	Interprestasi Parameter Loglinear.....	19
BAB IV	.....	21
METODOLOGI PENELITIAN	.....	21
4.1	Populasi Penelitian .....	21
4.2	Tempat dan Waktu Penelitian .....	21
4.3	Variabel Penelitian .....	21
4.4	Metode Pengumpulan Data .....	21
4.5	Alat Dan Metode Analisis Data.....	22
4.6	Langkah-Langkah Penelitian.....	22
BAB V	.....	25
HASIL DAN PEMBAHASAN	.....	25
5.1	Analisis Model Loglinear pada tabel kontigensi dua arah .....	25
5.1.1	Model Independent .....	25
5.1.2	Model Lengkap.....	27
5.1.3	Persamaan Likelihood .....	27
5.1.4	Uji Kecocokan Model.....	28
5.1.5	Partisi Chi-square untuk Membandingkan Model.....	29
5.2	Penerapan Model Loglinear.....	29
5.2.1	Menentukan Model Loglinear dan Loglinear penuh .....	30
5.2.2	Pendugaan Nilai Parameter .....	30
5.2.3	Uji Kecocokan Model.....	31
5.2.4	Interaksi dua variabel.....	33
5.2.5	Uji Parsial .....	35
5.3	Regresi Logistik.....	36
5.3.1	Regresi Logistik Multinomial.....	36

5.3.2	Uji Serentak .....	37
5.3.3	Uji Parsial .....	38
5.3.4	Uji Kesesuaian Model .....	38
5.4	Penerapan Model Logistik Multinomial.....	39
5.4.1	Pengujian Serentak .....	39
5.4.2	Pengujian Parsial .....	40
5.4.3	Pembentukan Model Logistik Multinomial.....	41
5.4.4	Uji kecocokan model .....	42
5.5	Menentukan nilai Prediksi Frekuensi Harapan.....	43
5.5.1	Loglinear.....	43
5.5.2	Logistik.....	44
5.5.3	Membandingkan Nilai AIC .....	47
BAB VI.....		48
PENUTUP.....		48
6.1	Kesimpulan.....	48
6.2	Saran .....	48



## DAFTAR TABEL

<b>Tabel 1.1.</b> Jumlah kasus tertinggi di wilayah Asia.....	2
<b>Tabel 2.1.</b> Tabel penelitian terdahulu .....	5
<b>Tabel 3.1.</b> Tabel kontigensi .....	9
<b>Tabel 3.2.</b> Tabel kontigensi dua arah.....	15
<b>Tabel 3.3.</b> Isi sel pada tabel silang dua arah Y dan X .....	20
<b>Tabel 4.1.</b> Variabel Penelitian .....	21
<b>Tabel 5.1.</b> Tabel persamaan <i>Likelihood</i> .....	27
<b>Tabel 5.2.</b> Tabel data kasus covid-19 .....	30
<b>Tabel 5.3.</b> Tabel estimasi nilai harapan .....	32
<b>Tabel 5.4.</b> Tabel uji interaksi dua variabel .....	33
<b>Tabel 5.5.</b> Tabel uji parsial .....	35
<b>Tabel 5.6.</b> <i>Likelihood</i> Ratio Test .....	39
<b>Tabel 5.7.</b> Hasil Estimasi Parameter.....	40
<b>Tabel 5.8.</b> Estimasi Parameter Loglinear .....	43
<b>Tabel 5.9.</b> Estimasi Parameter Logistik.....	44
<b>Tabel 5.10.</b> Tabel perbandingan model Loglinear dengan Logistik.....	47

## DAFTAR GAMBAR

<b>Gambar 4.1.</b> Diagram Alir Penelitian.....	24
---	----

## DAFTAR LAMPIRAN

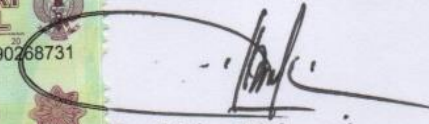
<b>Lampiran 1</b> Data Jumlah Kasus Covid-19 Tahun 2020 Menurut Negara Indonesia Dan Singapura Berdasarkan Jenis Status Sembuh, Meninggal Dan Dalam Perawatan .....	52
<b>Lampiran 2</b> Program Loglinear .....	53
<b>Lampiran 3</b> Program Logistik Multinomial .....	54
<b>Lampiran 4</b> Output Loglinear .....	55
<b>Lampiran 5</b> Output Logistik Multinomial .....	57
<b>Lampiran 6</b> Tabel <i>Chi-Square</i> .....	58

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali yang diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 10 Juli 2020



  
Ika Meiza Maharani

## ABSTRAK

### APLIKASI MODEL LOGLINEAR DAN REGRESI LOGISTIK

(Studi Kasus: Jumlah kasus Covid-19 pada tahun 2020)

Ika Meiza Maharani

Program Studi Statistika, Fakultas MIPA

Universitas Islam Indonesia

Tabel kontigensi merupakan tabel yang menunjukkan banyaknya individu yang diklasifikasikan menurut faktor-faktor tertentu. Tabel kontigensi dapat dimodelkan dengan menggunakan model loglinear dan logistik. Model loglinear digunakan untuk menyatakan suatu hubungan antara variabel-variabel kategorik (nominal atau ordinal). Untuk menganalisa model loglinear menggunakan uji *Chi-kuadrat* dan Uji *Rasio Likelihood*. Dalam model loglinear terdapat dua bentuk model, yaitu model loglinear penuh dan model loglinear tidak penuh, untuk model loglinear penuh maka perlu dilakukan pendugaan parameter dan pengujian interaksi antar variabelnya, sedangkan model loglinear tidak penuh hanya pendugaan parameter saja yang perlu dilakukan tidak ada pengujian interaksi antar variabel. Berdasarkan analisis data yang dilakukan model terbaik yang dipilih adalah model penuh yakni  $\ln m_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$ .

Sedangkan untuk model Logistik digunakan untuk menganalisis hubungan antara variabel prediktor dengan variabel respon yang bersifat kualitatif. Pada regresi Logistik model yang sesuai didapatkan setelah melakukan penaksiran parameter, uji signifikansi dan uji kecocokan model. Berdasarkan analisis data yang dilakukan didapatkan bahwa terdapat pengaruh antara Status ( $Y$ ) dengan variabel independen Negara ( $X$ ). Dan untuk uji kecocokan model, model dinyatakan layak atau tidak ada perbedaan yang nyata antara hasil dari observasi dengan hasil prediksi model.

**Kata Kunci:** *Loglinear, regresi Logistik, Chi-kuadrat, Rasio Likelihood.*

## ABSTRACT

### LOGLINEAR MODEL APPLICATION AND LOGISTIC REGRESSION

(Case Study : Covid-19 cases in 2020)

Ika Meiza Maharani

*Departement Of Statisticks, Faculty of Mathematics and Natural Sciences*

*University of Indonesia*

*Contingency table is a table that shows how many individuals are classified according to certain factors. Contingency tables can be modeled using loglinear and logistic models. The loglinear model is used to express a relationship between categorical variables (nominal or ordinal). To analyze the loglinear model using the Chi-Square test and Likelihood Ratio Test. In the loglinear model there are two forms of the model, namely the full loglinear model and the non-full loglinear model. For the full loglinear model, it is necessary to estimate parameters and test the interaction between the variables, whereas the loglinear model is not only estimating parameters that need to be done there is no interaction testing between variables. Based on data analysis, the best model chosen is the full model, namely  $\ln m_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$ .*

*Whereas the Logistics model is used to analyze the relationship between predictor variables and response variables that are qualitative in nature. In the logistic regression the model which is compiled is obtained after parameter estimation, significance test and model compatibility test. Based on the data analysis, it was found that there was an influence between Status (Y) and the independent variable Country (X). And for the model fit test, the model is declared to be feasible or there is no real difference between the results of the observations and the results of the model predictions.*

**Keywords:** *Loglinear, regresi Logisik, Chi-kuadrat, Rasio Likelihood.*

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Awal tahun 2020, dunia dikejutkan dengan mewabahnya *pneumonia* baru yang bermula dari Wuhan, Provinsi Hubei yang kemudian menyebar dengan cepat ke lebih dari 190 negara dan teritori. Wabah ini diberi nama *coronavirus disease* 2019 (COVID-19) yang disebabkan oleh *Severe Acute Respiratory Syndrome Coronavirus-2* (SARS-CoV-2). Penyebaran penyakit ini telah memberikan dampak luas secara sosial dan ekonomi. Masih banyak kontroversi seputar penyakit ini, termasuk dalam aspek penegakkan diagnosis, tata laksana, hingga pencegahan. (Susilo dkk, 2019). Virus ini dapat ditularkan dari manusia ke manusia dan telah menyebar secara luas di China dan lebih dari 190 negara dan teritori lainnya. Pada 12 Maret 2020, WHO mengumumkan COVID-19 sebagai pandemik. Hingga tanggal 29 Maret 2020, terdapat 634.835 kasus dan 33.106 jumlah kematian di seluruh dunia. Sementara di Indonesia sudah ditetapkan 1.528 kasus dengan positif COVID-19 dan 136 kasus kematian.

Informasi tentang virus ini masih sangat terbatas karena banyak hal masih dalam penelitian dan data epidemiologi akan sangat berkembang juga, untuk itu tinjauan ini merupakan tinjauan berdasarkan informasi terbatas yang dirangkum dengan tujuan untuk memberi informasi dan sangat mungkin akan terdapat perubahan kebijakan dan hal terkait lainnya sesuai perkembangan hasil penelitian, data epidemiologi dan kemajuan diagnosis dan terapi. Hal ini sesuai dengan kejadian penularan kepada petugas kesehatan yang merawat pasien COVID-19, disertai bukti lain penularan di luar Cina dari seorang yang datang dari Kota Shanghai, Cina ke Jerman dan diiringi penemuan hasil positif pada orang yang ditemui dalam kantor. Pada laporan kasus ini bahkan dikatakan penularan terjadi pada saat kasus indeks belum mengalami gejala (asimtomatik) atau masih dalam masa inkubasi. Laporan lain mendukung penularan antar manusia adalah laporan

9 kasus penularan langsung antar manusia di luar Cina dari kasus index ke orang kontak erat yang tidak memiliki riwayat perjalanan manapun.<sup>2,11</sup> Penularan ini terjadi umumnya melalui droplet dan kontak dengan virus kemudian virus dapat masuk ke dalam mukosa yang terbuka. Suatu analisis mencoba mengukur laju penularan berdasarkan masa inkubasi, gejala dan durasi antara gejala dengan pasien yang diisolasi. Analisis tersebut mendapatkan hasil penularan dari 1 pasien ke sekitar 3 orang di sekitarnya, tetapi kemungkinan penularan di masa inkubasi menyebabkan masa kontak pasien ke orang sekitar lebih lama sehingga risiko jumlah kontak tertular dari 1 pasien mungkin dapat lebih besar.

Berdasarkan data yang diterbitkan oleh WHO terdapat lima negara dengan jumlah kasus cukup tinggi untuk wilayah Asia Tenggara. Tabel 1.1 dibawah ini menunjukkan lima provinsi yang memiliki jumlah kasus tinggi untuk wilayah Asia Tenggara.

**Tabel 1.1.** Jumlah kasus tertinggi di wilayah Asia

	Total Cases	New Cases	Total Deaths	New Deaths	Total Recovered	Active Cases
<b>Indonesia</b>	41431	+1031	2276	+45	16243	22912
<b>Singapore</b>	41216	+247	26	0	31163	10027
<b>Philippines</b>	27238	+457	1108	+5	6820	19310
<b>Malaysia</b>	8151	+10	121	0	7873	521
<b>Thailand</b>	3135	0	58	0	2996	81
<b>Vietnam</b>	335	+1	0	0	325	10
<b>Myanmar</b>	262	0	6	0	179	77
<b>Brunei</b>	141	0	3	0	138	0
<b>Cambodia</b>	128	0	0	0	126	2
<b>Lao PDR</b>	19	0	0	0	19	0
<b>ASEAN</b>	122420	+1746	3598	+50	65882	52940

Berdasarkan hasil tabel 1.1 diatas menunjukkan lima Negara yang memiliki jumlah kasus terinfeksi visrus korona tertinggi di ASEAN diaman diperoleh bahwa Negara Indonesia menjadi negara pertama di ASEAN dengan jumlah kasus



infeksi virus corona terbanyak yaitu sebesar 41,431 kasus dan singapura merupakan Negara tertinggi kedua dengan jumlah kasus virus korana sebanyak 41,216 kasus. Akibat dari Tingkat jumlah kasus infeksi virus corona yang masih tinggi di Negara Indonesia dan Singapura memberikan indikasi bahwa diperlukan untuk dikaji dan dicermati secara komprehensif atas strategi, kebijakan dan program pengentasan kemiskinan yang telah diterapkan.

Dalam ilmu statistika banyak metode yang digunakan untuk menganalisis suatu kasus dalam sebuah penelitian, salah satunya yaitu metode analisis Loglinear dan Logistik. Model loglinear adalah salah satu kasus khusus pada *Generalized Linear Models* (GLM) untuk data yang berdistribusi Poisson. Model loglinear digunakan untuk menganalisis hubungan antara variabel respon yang bersifat kategori yang membentuk tabel kontingensi. Sedangkan model logistik atau model logit juga termasuk dalam model linear terampat *Generalized Linear Models* (GLM) yang digunakan untuk menganalisis hubungan antara peubah respon kategorik dengan satu atau beberapa peubah penjelas. (Agresti,1990)

Dalam sebuah penelitian banyak ditemukan keadaan dimana data yang dikumpulkan dapat dikategorikan menjadi dua atau lebih kategorik. Pada umumnya dalam menyajikan suatu data dapat dibuat dalam bentuk tabel. Tabel biasanya terdiri dari baris dan kolom yang menggambarkan suatu variabel dan frekuensinya. Dalam permasalahan ini data kategorik yang memiliki beberapa variabel disajikan dalam suatu tabel klasifikasi silang atau tabel kontigensi. Tabel kontigensi memiliki keuntungan yaitu lebih mudah dalam penyusunan suatu perhitungan, mudah menyajikan hasil analisis dan mempermudah dalam memahami rancangan yang kompleks. (Suryanto,1988)

Penerapan model loglinear dan logistik yang disusun dalam tabel kontingensi banyak ditemui pada kehidupan sehari-hari, salah satu contohnya seperti pada penelitian kali ini peneliti menggunakan data yang sedang hangat dibicarakan yaitu mengenai kasus Covid-19 berdasarkan negaranya pada tahun 2020. Variabel yang dimaksud yaitu Negara dan Jumlah Kasus Covid-19. Sehingga diketahui bahwa terdapat dua variabel yang akan diteliti, sehingga metode yang akan digunakan adalah model dua dimensi.

## **1.2 Rumusan Masalah**

Berdasarkan uraian latar belakang, maka didapatkan rumusan masalah yaitu bagaimana Aplikasi atau penggunaan model Loglinear dan Regresi Logistik pada tabel kontigensi dua dimensi dengan menggunakan kasus Covid-19 di Negara Indonesia dan Singapura?

## **1.3 Batasan Masalah**

Agar penyusunan laporan ini tidak terlalu meluas dan keluar dari pokok permasalahan, peneliti membatasi pokok permasalahan pada :

1. Bab 3 hanya membahas tentang definisi dan teorima yang berhubungan dengan model logistik dan loglinear, serta hubungan dari kedua model tersebut.
2. Data yang digunakan adalah data dengan tabel kontigensi 2 dimensi.
3. Data di olah menggunakan *software* R dan *SPSS* menggunakan metode Loglinear dan Logisik.

## **1.4 Tujuan Penelitian**

Berdasarkan pada rumusan masalah yang telah dipaparkan oleh peneliti, maka terdapat tujuan penelitian yang dapat disusun yaitu, melihat bagaimana penggunaan dari aplikasi model Loglinear dan Logistik dalam pemecahan suatu permasalahan dengan studi kasus yaitu kasus Covid-19 di Negara Indonesia dan Singapura menggunakan tabel kontigensi dua dimensi.

## **1.5 Manfaat Penelitian**

Berdasarkan tujuan yang ada, manfaat dari penelitian ini adalah untuk memperdalam tentang menganalisa tabel kontigensi khususnya dalam model logistik dan loglinear untuk tabel kontigensi dua dimensi serta penerapannya dalam menganalisis suatu masalah.

## BAB II

### TINJAUAN PUSTAKA

Setelah peneliti mengkaji terhadap penelitian terdahulu, ada beberapa kasus yang memiliki keterkaitan dengan penelitian yang peneliti lakukan. Peneliti akan meringkasnya dalam bentuk tabel di bawah ini:

**Tabel 2.1.** Tabel penelitian terdahulu

No	Nama peneliti	Judul Penelitian	Tahun	Isi
1.	Julio Adisantoso	<i>Model Loglinear dan Regresi Logistik</i>	2010	<ul style="list-style-type: none"><li>• Perbedaan Model Loglinear dan Regresi Logistik dengan menggunakan tabel 3 arah.</li><li>• Metode yang digunakan: Loglinear dan Logistik</li><li>• Software: SAS</li><li>• Perbedaan dengan penelitian ini : Studi kasus menggunakan data tabel 3 arah, aplikasinya menggunakan R.</li></ul>
2.	Mamik Lestiyorini	<i>Model Loglinear Multivariat Empat Dimensi</i>	2010	<ul style="list-style-type: none"><li>• Penggunaan Model Loglinear Multivariat dengan tabel 4 dimensi.</li><li>• Metode yang digunakan: Loglinear</li><li>• Software: R</li><li>• Perbedaan dengan penelitian ini : Pembahasan hanya model loglinear saja, untuk studi kasus menggunakan data tabel 4 dimensi.</li></ul>
3.	Try Azisah Nurman	<i>Analisis Data Kategori Dengan Loglinear Menggunakan Prinsip Hirarki</i>	2011	<ul style="list-style-type: none"><li>• Penggunaan model Loglinear dengan prinsip Hirarki</li><li>• Metode yang digunakan: Loglinear</li><li>• Perbedaan dengan penelitian ini : Pembahasan hanya model loglinear saja, Studi kasus menggunakan data tabel 3 arah, aplikasinya menggunakan</li></ul>

				R, metode menggunakan prinsip hirarki.
4.	Silvira Ayu Rosalia dan Sri Pingit Wulandari	<i>Analisis Model Loglinear Untuk Mengetahui Kecenderungan Perilaku Anak Jalanan Binaan Di Surabaya</i>	2010	<ul style="list-style-type: none"> <li>• Permodelan loglinear dengan data yang bersifat 2 kategorikal.</li> <li>• Metode yang digunakan Loglinear.</li> <li>• Perbedaan dengan penelitian ini : Pembahasan hanya model loglinear saja, aplikasinya menggunakan R.</li> </ul>
5.	Maryana	<i>Model Loglinear yang Terbaik untuk Analisis Data Kualitatif pada Tabel Kontingensi Tiga Arah</i>	2013	<ul style="list-style-type: none"> <li>• Penggunaan model Loglinear dengan tabel tiga arah</li> <li>• Metode yang digunakan: Loglinear</li> <li>• Software: SPSS</li> <li>• Perbedaan dengan penelitian ini : Pembahasan hanya model loglinear saja, Studi kasus menggunakan data tabel 3 arah, aplikasinya menggunakan R.</li> </ul>
6.	Dian Agustina dan Joko Purnomo	<i>Model Loglinear Faktor-faktor yang mempengaruhi Prilaku Merokok</i>	2014	<ul style="list-style-type: none"> <li>• Mendapatkan model terbaik interaksi dari faktor-faktor yang mempengaruhi waktu merokok</li> <li>• Hasil penelitian menggunakan model loglinear terdapat interaksi satu arah, dua arah, tiga arah, empat arah dan lima arah masuk kedalam model terbaik.</li> <li>• Software: SPSS</li> <li>• Perbedaan dengan penelitian ini : Pembahasan hanya model loglinear saja, studi kasus menggunakan data tabel 5 arah, aplikasinya menggunakan R.</li> </ul>
7.	Noor Lusty Putri Setiawati Dan Agung Priyono Utomo	<i>Model Regresi Logistik Untuk Melihat Pengaruh Faktor Demografis, Self Efficacy, Terhadap Perilaku Mencontek</i>	2017	<ul style="list-style-type: none"> <li>• Penggunaan model logistik untuk menentukan pengaruh beberapa faktor.</li> <li>• Metode yang digunakan: Regresi Logistik</li> <li>• Software : SPSS</li> <li>• Perbedaan dengan penelitian ini : Pembahasan hanya regresi logistik saja,</li> </ul>

				aplikasinya menggunakan R.
8.	Yumira Adriani Tampil, Hanny Komalig, dan Yohanis Langi.	<i>Analisis Regresi Logistik Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa FMIPA Universitas Sam Ratulangi Manado</i>	2017	<ul style="list-style-type: none"> <li>• Penggunaan Regresi Logistik Biner dalam menentukan faktor-faktor yang mempengaruhi.</li> <li>• Metode yang digunakan: Regresi Logistik</li> <li>• Software : SPSS</li> <li>• Perbedaan dengan penelitian ini : Pembahasan hanya regresi logistik saja, Studi kasus menggunakan data IPK mahasiswa, aplikasinya menggunakan R.</li> </ul>
9.	Rio A. Kasengkang, Sientje Nangoy, dan Jacky Sumarauw.	<i>Analisis Logistik (Studi Kasus Pada Pt. Remenia Satori Tepas-Kota Manado)</i>	2016	<ul style="list-style-type: none"> <li>• Konsep penggunaan metode analisis Logistik dalam penerapan suatu kasus.</li> <li>• Metode yang digunakan: Regresi Logistik</li> <li>• Perbedaan dengan penelitian ini : Pembahasan hanya regresi logistik saja, aplikasinya menggunakan R</li> </ul>
10.	Riska Habriel Ruslie dan Darmadi	<i>Analisis Regresi Logistik Untuk Faktor-Faktor Yang Mempengaruhi Status Gizi Remaja</i>	2012	<ul style="list-style-type: none"> <li>• Mengetahui prevalensi status gizi mahasiswa dan faktor-faktor apa saja yang mempengaruhi status gizi pada remaja,</li> <li>• Metode yang digunakan: Regresi Logistik</li> <li>• Perbedaan dengan penelitian ini : Pembahasan hanya regresi logistik saja,</li> </ul>

Berdasarkan beberapa referensi penelitian pada tabel 2.1, beberapa pengembangan Penelitian yang dilakukan oleh Maharani (2020) yang berjudul “Aplikasi model Loglinear dan Regresi Logistik” dengan variabel yang diteliti yaitu Negara dan Kasus Covid. Adapun tujuan dari penelitian ini yaitu untuk mengetahui bagaimana penggunaan dari aplikasi model Loglinear dan Logistik dalam pemecahan suatu permasalahan dengan studi kasus yaitu kasus Covid-19 di Negara Indonesia dan Singapura menggunakan tabel kontigensi dua dimensi. Analisis yang digunakan adalah analisis model loglinear dan analisis Regresi Logistik Multinomial.

## BAB III

### LANDASAN TEORI

#### 3.1 Coronavirus Disease (Covid-19)

*Severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2) yang lebih dikenal dengan nama virus Corona adalah jenis baru dari coronavirus yang menular ke manusia. Infeksi virus Corona disebut COVID-19 (*Corona Virus Disease 2019*) dan pertama kali ditemukan di kota Wuhan, China pada akhir Desember 2019. Virus ini menular dengan sangat cepat dan telah menyebar ke hampir semua negara, termasuk Indonesia, hanya dalam waktu beberapa bulan.

Coronavirus adalah kumpulan virus yang bisa menginfeksi sistem pernapasan. Pada banyak kasus, virus ini hanya menyebabkan infeksi pernapasan ringan, seperti flu. Namun, virus ini juga bisa menyebabkan infeksi pernapasan berat, seperti infeksi paru-paru (pneumonia). Selain virus SARS-CoV-2 atau virus Corona, virus yang juga termasuk dalam kelompok ini adalah virus penyebab *Severe Acute Respiratory Syndrome* (SARS) dan virus penyebab *Middle-East Respiratory Syndrome* (MERS). Meski disebabkan oleh virus dari kelompok yang sama, yaitu coronavirus, COVID-19 memiliki beberapa perbedaan dengan SARS dan MERS, antara lain dalam hal kecepatan penyebaran dan keparahan gejala.

Gejala awal infeksi virus Corona atau COVID-19 bisa menyerupai gejala flu, yaitu demam, pilek, batuk kering, sakit tenggorokan, dan sakit kepala. Setelah itu, gejala dapat hilang dan sembuh atau malah memberat. Penderita dengan gejala yang berat bisa mengalami demam tinggi, batuk berdahak bahkan berdarah, sesak napas, dan nyeri dada. (Alodokter, 2020)

#### 3.2 Tabel Kontigensi Dua Dimensi

Secara umum, tabel kontingensi dua dimensi dapat disajikan dalam bentuk tabel I x J. Tabel I x J terdapat dua variabel yaitu variabel A dan variabel B. Dalam tabel ini mempunyai I baris yang menyatakan kategori dari variabel A dan

J kolom yang menyatakan kategori dari variabel B. Terdapat IJ sel dalam tabel yang berisi frekuensi pengamatan yang terjadi dari kombinasi kedua kategori variabel sehingga diperoleh data berkategori dalam bentuk kontingensi dua dimensi berukuran I x J (Hapsari, 2011). Berikut tabel kontingensi untuk dua dimensi.

**Tabel 3.1.** Tabel kontingensi

		B				Total
		B <sub>1</sub>	B <sub>2</sub>	...	B <sub>j</sub>	
A	A <sub>1</sub>	$n_{11}$	$n_{12}$	...	$n_{1j}$	$n_{1.}$
	A <sub>2</sub>	$n_{21}$	$n_{22}$	...	$n_{2j}$	$n_{2.}$
	⋮	⋮	⋮	...	⋮	⋮
	A <sub>i</sub>	$n_{i1}$	$n_{i2}$	...	$n_{ij}$	$n_{i.}$
	Total	$n_{.1}$	$n_{.2}$	...	$n_{.j}$	$n_{..}$

Keterangan:

$n_{ij}$  = Frekuensi pengamatan pada baris ke-i dan kolom ke-j

$n_{i.}$  = Total marginal pada variabel baris ke-i

$n_{.j}$  = Total marginal pada variabel kolom ke-j

$n_{..}$  = Total frekuensi pengamatan

### 3.3 Regresi Logistik

Regresi logistik merupakan permodelan statistik dengan pendekatan seperti regresi linear, namun hanya saja pada regresi logistik memiliki variabel terikat yang berskala dikotomi, yaitu skala data nominal dengan dua kategori yang dinyatakan sebagai “sukses” dan “gagal”.

#### 3.3.1 Model Regresi Logistik

Model regresi logistik sederhana yaitu model regresi logistik untuk satu variabel prediktor X dengan variabel respon Y yang bersifat dikotomi. Nilai variabel Y = 1 menyatakan adanya suatu karakteristik dan Y = 0 menyatakan tidak adanya suatu karakteristik. Menurut (Hosmer, 1989) model regresi logistik

yang dipengaruhi oleh  $p$  variabel prediktor dapat dinyatakan sebagai nilai harapan dari  $Y$  dengan diberikan nilai  $x$

$$E(Y|x) = \frac{e^{(\beta_0 + \sum_{k=1}^p \beta_k x_k)}}{1 + e^{(\beta_0 + \sum_{k=1}^p \beta_k x_k)}} \quad (3.1)$$

dengan  $0 \leq E(Y|x) \leq 1$  dan  $Y$  mempunyai nilai 0 atau 1. Nilai  $E(Y|x)$  merupakan probabilitas sukses, sehingga dapat dinyatakan dengan  $p(x)$ , sehingga persamaan (3.1) menjadi

$$p(x) = \frac{e^{(\beta_0 + \sum_{k=1}^p \beta_k x_k)}}{1 + e^{(\beta_0 + \sum_{k=1}^p \beta_k x_k)}} \quad (3.2)$$

dengan  $\beta_k$  menyatakan parameter-parameter regresi,  $x_k$  adalah pengamatan variabel prediktor ke- $k$  dari sejumlah  $p$  variabel prediktor.

Transformasi logit diterapkan pada model regresi logistik,

$$\text{Logit}(p(x)) = g(x) = \ln \left[ \frac{p(x)}{1 - p(x)} \right] = \beta_0 + \sum_{k=1}^p \beta_k x_k \quad (3.3)$$

Transformasi logit bertujuan untuk membuat fungsi linear dari parameter-parameternya. Fungsi  $g(x)$  linear terhadap parameter dan memiliki range  $(-\infty, \infty)$ , tergantung dari range variabel prediktor  $X$ .

### 3.3.2 Penaksiran Parameter Model Logistik

Metode penaksiran parameter yang biasa digunakan dalam regresi logistik adalah metode maksimum *Likelihood*. Setiap observasi untuk model regresi logistik adalah variabel random dari distribusi Bernoulli.

Menurut Hosmer dan Lemeshow (1989), fungsi *Likelihood* distribusi Bernoulli untuk  $n$  sampel independen adalah

$$l(\beta) = \prod_{i=1}^n p(x_i)^{Y_i} (1 - p(x_i))^{1-Y_i} \quad (3.4)$$

dan log-*Likelihood* atau logaritma natural fungsi probabilitas bersamanya adalah



$$\begin{aligned}
L(\beta) &= \ln l(\beta) \\
&= \ln \prod_{i=1}^n p(x_i)^{Y_i} (1 - p(x_i))^{1-Y_i} \\
&= \sum_{i=1}^n Y_i(\beta_0 + \beta_i x_i) - \sum_{i=1}^n \ln[1 + \exp(\beta_0 + \beta_i x_i)]
\end{aligned} \tag{3.5}$$

Taksiran parameter  $\beta_k$  ,diperoleh dengan mendiferensialkan fungsi *log-Likelihood* terhadap  $\beta_k$  dengan  $k = 0, 1$ . Nilai maksimum diperoleh bila hasil diferensial fungsi *log-Likelihood* bernilai nol. Diperlukan metode iterasi untuk mendapatkan taksiran pada metode maksimum *Likelihood* karena tidak bisa diperoleh taksiran parameter dari pendefrensialan fungsi *log-Likelihood*.

### 3.3.3 Uji Signifikansi Parameter Logistik

Uji signifikansi parameter dari variabel prediktor dilakukan untuk mengetahui apakah taksiran parameter yang diperoleh berpengaruh secara signifikan terhadap model atau tidak, dan seberapa besar pengaruh masing-masing parameter tersebut terhadap model.

Uji signifikansi terdiri dari dua tahap yaitu uji signifikansi parameter model secara bersama dan uji signifikansi parameter model secara terpisah. Uji signifikansi parameter model secara bersama dilakukan dengan uji rasio *Likelihood G* adalah fungsi dari  $L_0$  dan  $L_1$  yang berdistribusi  $\chi^2$  (*Chi-square*) dengan derajat bebas  $p$  (banyaknya variabel prediktor yang ada dalam model) yang didefinisikan sebagai

$$G = -2 (L_0 - L_1) \tag{3.6}$$

dengan  $L_0$  adalah *log-Likelihood* dari model tanpa variabel prediktor, sedangkan  $L_1$  adalah *log-Likelihood* dari model dengan  $p$  variabel prediktor. Nilai *log-Likelihood* dihitung berdasarkan persamaan (3.5). Hipotesis nolnya adalah  $\beta_k = 0$ , untuk semua  $k = 1, 2, \dots, p$ , yang berarti bahwa semua variabel prediktor tidak signifikan terhadap model. Hipotesis nol ditolak jika  $G > \chi^2_{(\alpha;p)}$ . Jika nilai prediksi ketika variabel prediktor di dalam model lebih baik daripada ketika

variabel tersebut tidak disertakan dalam model, maka dapat dikatakan variabel signifikan di dalam model.

Uji Wald Chi-square digunakan untuk menguji signifikansi parameter model secara terpisah. Uji Wald Chi-square didefinisikan dengan

$$W_k = \left[ \frac{\hat{\beta}_k}{\hat{SE}(\hat{\beta}_k)} \right]^2, \quad (3.7)$$

Dengan  $k = 1, 2, \dots, p$ .

Statistik uji  $W_k$  mendekati distribusi *Chi-square* dengan derajat bebas 1. Hipotesis nolnya adalah  $\beta_k = 0$ , untuk setiap  $k = 1, 2, \dots, p$ , yang berarti bahwa variabel prediktor ke- $k$  tidak signifikan terhadap model. Hipotesis nol ditolak jika  $W_k > \chi^2_{(\alpha; 1)}$

### 3.3.4 Uji Kecocokan Model Logistik

Uji kecocokan model digunakan untuk mengevaluasi cocok tidaknya model dengan data, nilai observasi yang diperoleh sama atau mendekati dengan yang diharapkan dalam model. Cocok tidaknya model regresi logistik pada skripsi ini dinilai dengan menggunakan uji Hosmer dan Lemeshow karena terdapat variabel prediktor yang bersifat kontinu, yaitu pendapatan per kapita. Variabel tersebut memungkinkan terjadinya pola kovariat yang beragam, sehingga uji Hosmer dan Lemeshow lebih tepat untuk diterapkan. Uji Hosmer dan Lemeshow dapat digunakan saat pola kovariat yang sama dari variabel prediktor muncul dalam observasi atau tidak. Pola kovariat merupakan kejadian dari nilai-nilai variabel prediktor. Jika semua pola kovariat dari variabel prediktor merupakan kejadian unik, maka jumlah pola kovariatnya sama dengan jumlah sampel ( $n$ ). Jika uji Hosmer dan Lemeshow dipenuhi maka model mampu memprediksi nilai observasinya atau dapat dikatakan model dapat diterima karena sesuai dengan data observasinya.

Uji Hosmer dan Lemeshow yang ditulis dengan uji  $\hat{C}$ , dihitung berdasarkan taksiran probabilitas, (Hosmer, 1989). Pada uji ini sampel dimasukkan ke sejumlah  $g$  kelompok dengan tiap-tiap kelompok memuat  $n/10$  sampel pengamatan, dengan  $n$  adalah jumlah sampel. Jumlah kelompok sekitar 10.

Idealnya, kelompok pertama memuat  $n'_1 = n/10$  sampel yang memiliki taksiran probabilitas sukses terkecil yang diperoleh dari model taksiran. Kelompok kedua memuat  $n'_2 = n/10$  sampel yang memiliki taksiran probabilitas sukses terkecil kedua, dan seterusnya, (Liu, 2007).

Statistik uji Hosmer dan Lemeshow  $\hat{C}$  yang dihitung berdasarkan nilai  $y = 1$  dirumuskan

$$\hat{C} = \sum_{r=1}^g \frac{(O_r - n'_r \bar{p}_{1r})^2}{n'_r \bar{p}_{1r} (1 - \bar{p}_{1r})} \quad (3.8)$$

dengan  $\bar{p}_{1r}$  menyatakan rata-rata taksiran probabilitas sukses kelompok ke- $r$ ,  $O_r$  adalah jumlah sampel kejadian sukses dalam kelompok ke- $r$ ,  $n'_r$  adalah total sampel kelompok ke- $r$ ,  $\sum_{r=1}^g n'_r = n$  dan, dengan  $r = 1, 2, \dots, g$ . Statistik uji  $\hat{C}$  mendekati distribusi *Chi-Square* dengan derajat bebas  $g-2$ , (Hosmer dan Lemeshow, 1989). Hipotesis nol menyatakan bahwa model cocok dengan data. Hipotesis nol ditolak jika  $\hat{C} > \chi^2_{(\alpha; g-2)}$

### 3.3.5 Interpretasi Parameter Logistik

Pada pemodelan regresi logistik, interpretasi parameter bertujuan untuk mengetahui arti dari nilai taksiran parameter pada variabel prediktor. Terdapat dua jenis variabel prediktor, yaitu variabel yang bersifat kategorik dan variabel kontinu.

Cara yang digunakan untuk menginterpretasikan parameter regresi logistik dari variabel kategorik adalah dengan rasio odds, (Hosmer, 1989).

*Odds* adalah perbandingan probabilitas kejadian sukses dengan kejadian tidak sukses dalam suatu kategori. *Odds* untuk  $x = 1$  dan  $x = 0$  secara berturut-turut adalah

$$\frac{p(1)}{1 - p(1)} \text{ dan } \frac{p(0)}{1 - p(0)}$$

Rasio Odds merupakan perbandingan nilai odds untuk kategori  $x = 1$  terhadap *odds* untuk kategori  $x = 0$ , dalam variabel prediktor yang sama dengan menganggap variabel prediktor lainnya konstan.

Rasio Odds dinyatakan dengan  $\psi$  dan dituliskan sebagai

$$\psi = \frac{\left(\frac{p(1)}{1-p(1)}\right)}{\frac{p(0)}{1-p(0)}} \quad (3.9)$$

Kehadiran variabel  $x$  dengan nilai 1 akan memberikan nilai  $\psi$  kali dibanding  $x$  dengan nilai 0 untuk menghasilkan kejadian sukses  $Y = 1$ .

Cara yang digunakan untuk menginterpretasi parameter regresi logistik dari variabel kontinu adalah dengan mengasumsikan fungsi logit linear terhadap variabel prediktor. Dimisalkan variabel prediktornya kontinu, dan fungsi  $g(x) = \beta_0 + \beta_1 x_1$  interpretasi dari  $\beta_1$  memiliki sifat sama dengan parameter pada regresi linear. Setiap kenaikan satu unit  $x_1$ , nilai  $g(x)$  naik sebesar  $\beta_1$ , dapat dinyatakan dengan  $\beta_1 = g(x + 1) - g(x)$  untuk setiap nilai  $x$ .

### 3.4 Model Loglinear

Model loglinear adalah suatu model untuk menyatakan hubungan antara variabel dengan data yang bersifat kualitatif (skala nominal atau ordinal). Model loglinear merupakan salah satu khusus dari *Generalized Linear Model* (GLM). Analisis dengan model loglinear dilakukan untuk mempelajari pola asosiasi antara sekelompok variabel, selain itu juga digunakan untuk memperkirakan banyaknya observasi yang diharapkan dalam tiap sel tabel kontigensi. Selanjutnya berdasarkan nilai harapan itu dapat dihitung beberapa statistik penting, seperti populasi, statistik rasio kecenderungan.

#### 3.4.1 Model Loglinear dalam tabel kontigensi dua arah

Permodelan loglinear merupakan permodelan frekuensi observasi dalam tabulasi silang. Permisalan data disajikan dalam tabel kontigensi dua arah berukuran  $b \times k$ . Dimana  $b$  adalah banyaknya baris dan  $k$  adalah banyaknya kolom. Yang disajikan seperti pada tabel (3.2)

**Tabel 3.2.** Tabel kontigensi dua arah

Faktor A	Faktor Y				Jumlah
	Y <sub>1</sub>	Y <sub>2</sub>	...	Y <sub>j</sub>	
X <sub>1</sub>	n <sub>11</sub> π <sub>11</sub>	n <sub>12</sub> π <sub>12</sub>	...	n <sub>1j</sub> π <sub>1j</sub>	n <sub>1+</sub> π <sub>1+</sub>
X <sub>2</sub>	n <sub>21</sub> π <sub>21</sub>	n <sub>22</sub> π <sub>22</sub>	...	n <sub>2j</sub> π <sub>2j</sub>	n <sub>2+</sub> π <sub>2+</sub>
⋮	...	...	...	...	...
X <sub>i</sub>	n <sub>i1</sub> π <sub>i1</sub>	n <sub>i2</sub> π <sub>i2</sub>	...	n <sub>ij</sub> π <sub>ij</sub>	n <sub>i+</sub> π <sub>i+</sub>
Jumlah	n <sub>+1</sub> π <sub>+1</sub>	n <sub>+2</sub> π <sub>+2</sub>	...	n <sub>+j</sub> π <sub>+j</sub>	n <sub>++</sub> π <sub>++</sub>

Keterangan:

$n_{ij}$  = Frekuensi pengamatan pada baris ke-i dan kolom ke-j

$\pi_{ij}$  = Peluang pengamatan pada baris ke-i dan kolom ke-j

Jika kedua faktor ini independen, maka peluang pengamatan pada baris ke-i dan kolom ke-j yaitu  $\pi_{ij}$  sama dengan perkalian peluang margin baris ke-i ( $\pi_{i+}$ ) dengan peluang margin kolom ke-j ( $\pi_{+j}$ ).

$$\pi_{ij} = \pi_{i+} \cdot \pi_{+j}$$

dimana frekuensi nilai harapannya adalah sebagai berikut.

$$\hat{m}_{ij} = n_{++} \pi_{ij} = n_{++} \pi_{i+} \pi_{+j} \quad (3.10)$$

Jika persamaan (3.10) dinyatakan dalam bentuk logaritma, maka didapatkan:

$$\log m_{ij} = \log n_{++} + \log \pi_{i+} + \log \pi_{+j}$$

Bila dijumlahkan untuk semua i (baris) maka:

$$a. \sum_{i=1}^I \log m_{ij} = I \log n_{++} + \sum_{i=1}^I \pi_{i+} + I \log \pi_{+j}$$

Dan bila dijumlahkan untuk semua j (kolom), maka model menjadi :

$$b. \sum_{j=1}^J \log m_{ij} = J \log n_{++} + J \log \pi_{i+} + \sum_{j=1}^J \log \pi_{+j}$$

Sehingga bila dijumlahkan untuk semua i dan j, didapat :

$$c. \sum_{i=1}^I \sum_{j=1}^J \log m_{ij} = IJ \log n_{++} + J \sum_{i=1}^I \log \pi_{i+} + I \sum_{j=1}^J \log \pi_{+j}$$

Berdasarkan persamaan (a), (b) dan (c) dengan mengambil

$$\mu = \log n_{++} + \left[ \sum_{i=1}^I \frac{\log \pi_{i+}}{I} \right] + \left[ \sum_{j=1}^J \frac{\log \pi_{+j}}{J} \right]$$

$$\lambda_i^x = \log \pi_{i+} - \left[ \sum_{i=1}^I \frac{\log \pi_{i+}}{I} \right]$$

$$\lambda_j^y = \log \pi_{+j} - \left[ \sum_{j=1}^J \frac{\log \pi_{+j}}{J} \right]$$

Akan diperoleh persamaan

$$\log m_{ij} = \mu + \lambda_i^x + \lambda_j^y \quad (3.11)$$

Persamaan (3.11) inilah yang disebut dengan model Loglinear Independen pada tabel kontingensi dua dimensi (Agresti A. , 1984). Dalam model tersebut  $\mu$  menunjukkan efek rata-rata secara umum,  $\lambda_i^x$  menunjukkan efek utama kategori ke-i variabel X,  $\lambda_j^y$  menunjukkan efek utama kategori ke-j variabel Y. Dimana juga berlaku  $\sum_{i=1}^I \lambda_i^x = \sum_{j=1}^J \lambda_j^y = 0$ .

Jika ada dependensi antara kedua variabel, dengan nilai  $m_{ij} > 0$  dan dimisalkan

$$\eta_{ij} = \log m_{ij}$$

$$\eta_{i+} = \sum_{j=1}^J \frac{\eta_{ij}}{J}$$

$$\eta_{+j} = \sum_{i=1}^I \frac{\eta_{ij}}{I}$$

$$\mu = \eta_{++} = \sum_{i=1}^I \sum_{j=1}^J \frac{\eta_{ij}}{IJ}$$

Serta jika ditetapkan

$$\lambda_i^x = \eta_{i+} - \eta_{++}$$

$$\lambda_j^y = \eta_{+j} - \eta_{++}$$

$$\lambda_{ij}^{xy} = \eta_{ij+} - \eta_{i+} - \eta_{+j} + \eta_{++}$$

Maka model persamaan menjadi sebagai berikut.

$$\log m_{ij} = \mu + \lambda_i^x + \lambda_j^y + \lambda_{ij}^{xy} \quad (3.12)$$

Persamaan (3.12) disebut dengan model jenuh. Selanjutnya dicari nilai dari derajat bebasnya (df). Derajat bebas adalah banyaknya sel dikurangi dengan banyaknya parameter yang diestimasi. Untuk model independen(3.11), merupakan kasus khusus dari model jenuh (3.12) dimana  $\lambda_{ij}^{xy} = 0$ . Jumlah parameter yang diestimasi  $= I + (I - 1) + (J - 1)$ . Sehingga untuk model independen, mempunyai derajat bebas

$$\begin{aligned} df &= (IJ - 1) - [(I - 1) + (J - 1)] \\ &= IJ - 1 - J + 1 \\ &= (I - 1)(J - 1) \end{aligned}$$

### 3.4.2 Penaksiran Parameter Model Loglinear

Untuk menaksir parameter-parameternya dapat digunakan metode *Maximum Likelihood Estimator* (MLE). *Maksimum Likelihood* adalah teknik yang digunakan dalam penaksiran nilai parameter bila distribusi populasi diketahui.

*Definisi 2. 1*

Misalkan  $X_1, X_2, \dots, X_n$  peubah acak dengan fungsi distribusi  $f(x_1, x_2, \dots, x_n | \theta)$  dengan  $\theta \in \Theta$  yang tidak diketahui, maka fungsi *Likelihood* ialah:

$$L(\theta) = \begin{cases} f(x_1, x_2, \dots, x_n | \theta) & \text{jika } F \text{ mempunyai fungsi padat } f \\ P(x_1, x_2, \dots, x_n | \theta) & \text{jika } F \text{ mempunyai fungsi padat } P \end{cases}$$

Untuk setiap  $\hat{\theta} = \hat{\theta}_n(X_1, X_2, \dots, X_n) \in \Theta$  sehingga  $L\hat{\theta} = \sup \{L(\theta) : \theta \in \Theta\}$  disebut *maximum Likelihood estimation* (Misbahussurur, 2009)

Langkah-langkah estimasi maksimum *Likelihood*:

- Menentukan fungsi distribusi.
- Menentukan fungsi *Likelihood* dari fungsi distribusi.
- Menentukan fungsi maksimum *Likelihood* (*log Likelihood*) dari fungsi distribusi.
- Menentukan penduga parameter-parameter dengan memaksimumkan fungsi maksimum *Likelihood* dari fungsi distribusi yang telah ditentukan.

### 3.4.3 Uji Kecocokan Model Loglinear

Uji kecocokan model merupakan salah satu yang terpenting dalam analisa data kualitatif yang digunakan untuk membandingkan hasil observasi dengan nilai harapannya (Sudjana, 1990). Hasil perbandingan tersebut dapat diketahui model loglinear yang cocok untuk analisa data hasil pengamatan dengan hipotesisnya adalah  $H_0$  menyatakan bahwa model loglinear cocok, sedangkan hipotesis alternatif  $H_1$  menyatakan bahwa model loglinear tidak cocok. Uji hipotesis ini menggunakan statistik uji sebagai berikut:

1. Statistik Pearson atau Chi-kuadrat

Statistik Chi-kuadrat untuk table kontingensi dua arah adalah

$$X^2 = \sum_{i=1}^I \sum_{j=1}^J \left( \frac{n_{ij} - \hat{m}_{ij}}{\hat{m}_{ij}} \right)^2 \quad (3.13)$$

Keterangan :

$X^2$  = Statistik uji Chi-kuadrat

$n_{ij}$  = Jumlah observasi

$\hat{m}_{ij}$  = Nilai harapan

2. Statistik perbandingan *Likelihood* Chi-kuadrat

Statistik perbandingan *Likelihood* Chi-kuadrat untuk tabel kontingensi tiga arah adalah

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log \left( \frac{n_{ij}}{\hat{m}_{ij}} \right) \quad (3.14)$$

Keterangan :

$G^2$  = Statistik uji perbandingan

$n_{ij}$  = Jumlah observasi

$\hat{m}_{ij}$  = Nilai harapan

Bila model yang ditentukan jumlah observasinya besar maka baik  $X^2$  dan  $G^2$  mendekati distribusi Chi-kuadrat dengan derajat kebebasan sama dengan jumlah sel dikurangi jumlah parameter-parameter bebas yang terdapat dalam model. Kriteria pengujian hipotesisnya adalah tolak  $H_0$  apabila  $\chi^2$  hitung lebih



besar dari  $\chi^2$  tabel atau probabilitas (*p-value*) lebih kecil dari taraf *signifikansi* ( $\alpha$ ).

#### 3.4.4 Penyeleksi Model Terbaik

Menyeleksi model terbaik dengan metode *elimination backward* yang berdasarkan prinsip hirarki dimulai dari model jenuh (saturated) dan secara berurutan mengeliminasi model. Pada setiap tahap, metode ini mengeliminasi model yang mempunyai pengaruh atau efek yang paling kecil. Proses seleksi berhenti ketika beberapa pengeliminasian yang dilakukan telah mendapatkan model yang sesuai dan lebih sederhana. Langkah-langkah yang dilakukan adalah (Wulandari, 2011)

- a. Anggap model (1) adalah model XY sebagai model terbaik.
- b. Keluarkan efek interaksi dua variabel sehingga modelnya (X,Y) yang disebut dengan model (2).
- c. Bandingkan model (1) dengan model (2) dengan hipotesis sebagai berikut:  
 $H_0$  : model (2) adalah model yang terbaik  
 $H_1$  : model (1) bukan model yang terbaik

Statistik uji yang digunakan:  $G^2$  (*Likelihood Ratio*) Kriteria penolakan:  $G^2 > \chi^2$  tabel maka tolak  $H_0$ .

- i. Jika  $H_0$  ditolak, maka dinyatakan bahwa model (1) adalah yang terbaik. Tetapi jika gagal tolak  $H_0$ , maka bandingkan kembali model (2) dengan model (1) dengan mengeluarkan salah satu indeks faktor.
- ii. Untuk menentukan indeks faktor mana yang dikeluarkan terlebih dahulu maka dipilih nilai  $G^2$  terkecil.

#### 3.4.5 Interpretasi Parameter Loglinear

Cara yang digunakan untuk menginterpretasikan parameter loglinear sama dengan logistik yaitu menggunakan rasio odds.

Dengan demikian, rasio odd untuk tabel silang dua arah adalah (Nugraha, 2017):

$$\frac{\mu_{00}/\mu_{01}}{\mu_{10}/\mu_{11}} = \frac{\mu_{00}/\mu_{11}}{\mu_{11}/\mu_{10}} = \frac{e^{\mu} \cdot e^{\mu + \alpha_1^X + \beta_1^Y + (\alpha\beta)_{11}^{XY}}}{e^{\mu + \beta_1^Y} \cdot e^{\mu + \alpha_1^X}} = e^{(\alpha\beta)_{11}^{XY}} \quad (3.15)$$

Maka didapatkan tabel silang dua arah Y dan X adalah:

**Tabel 3.3.** Isi sel pada tabel silang dua arah Y dan X

		Y	
		0	1
X	0	$e^{\mu}$	$e^{\mu+\beta_1^Y}$
	1	$e^{\mu+\alpha_1^X}$	$e^{\mu+\alpha_1^X+\beta_1^Y+(\alpha\beta)_{11}^{XY}}$

Berdasarkan tabel (3.3) dapat diketahui bahwa  $(\alpha\beta)_{11}^{XY}$  merupakan log dari rasio odds. Jika  $(\alpha\beta)_{11}^{XY} = 0$ , maka tidak ada interaksi antara X dan Y sehingga menjadi model *additive* sempurna, yaitu.

$$\log \mu_{ij} = \mu + \alpha_i^x + \beta_j^y \quad (3.16)$$

## BAB IV

### METODOLOGI PENELITIAN

#### 4.1 Populasi Penelitian

Populasi dalam penelitian ini adalah seluruh penduduk Indonesia dan Singapura. Sedangkan untuk sampel adalah penduduk yang terdampak Covid-19 hingga bulan April 2020 yang terdiri dari 23007 jiwa.

#### 4.2 Tempat dan Waktu Penelitian

Penelitian ini dilakukan di rumah. Penelitian dilakukan pada bulan April 2020

#### 4.3 Variabel Penelitian

Dalam penelitian ini terdapat dua jenis variabel yang menjadi obyek penelitian, peneliti mengambil negara Indonesia dan Singapura yaitu:

**Tabel 4.1.** Variabel Penelitian

Variabel	Keterangan	Kategori
<i>Independen variable</i> (variabel X)	Negara	0. Indonesia
		1. Singapura
<i>Dependent variable</i> (variabel Y)	Status kasus Covid-19	1. Sembuh
		2. Meninggal
		3. Dalam perawatan

#### 4.4 Metode Pengumpulan Data

Data-data yang digunakan dalam penelitian ini adalah data sekunder yang didapat dari *website* <https://www.worldometers.info/coronavirus/#countries>. pada bulan April 2020. Data yang digunakan dapat dilihat pada variabel yang digunakan dalam penelitian. Setiap variabel yang digunakan merupakan data yang terkumpul pada tahun 2020.

#### 4.5 Alat Dan Metode Analisis Data

Dalam menganalisis data-data yang terdapat pada penelitian ini, peneliti akan menggunakan beberapa bantuan *software* yang akan digunakan untuk membahas dan menganalisa data penelitian. *Software* yang digunakan oleh peneliti adalah *software R* serta *software* pendukung lainnya. Sedangkan untuk metode analisis data nya menggunakan metode Loglinear dan Logistik Multinomial.

#### 4.6 Langkah-Langkah Penelitian

Berikut ini langkah-langkah yang dilakukan, yaitu :Metode analisis untuk model loglinear dengan langkah-langkah sebagai berikut.

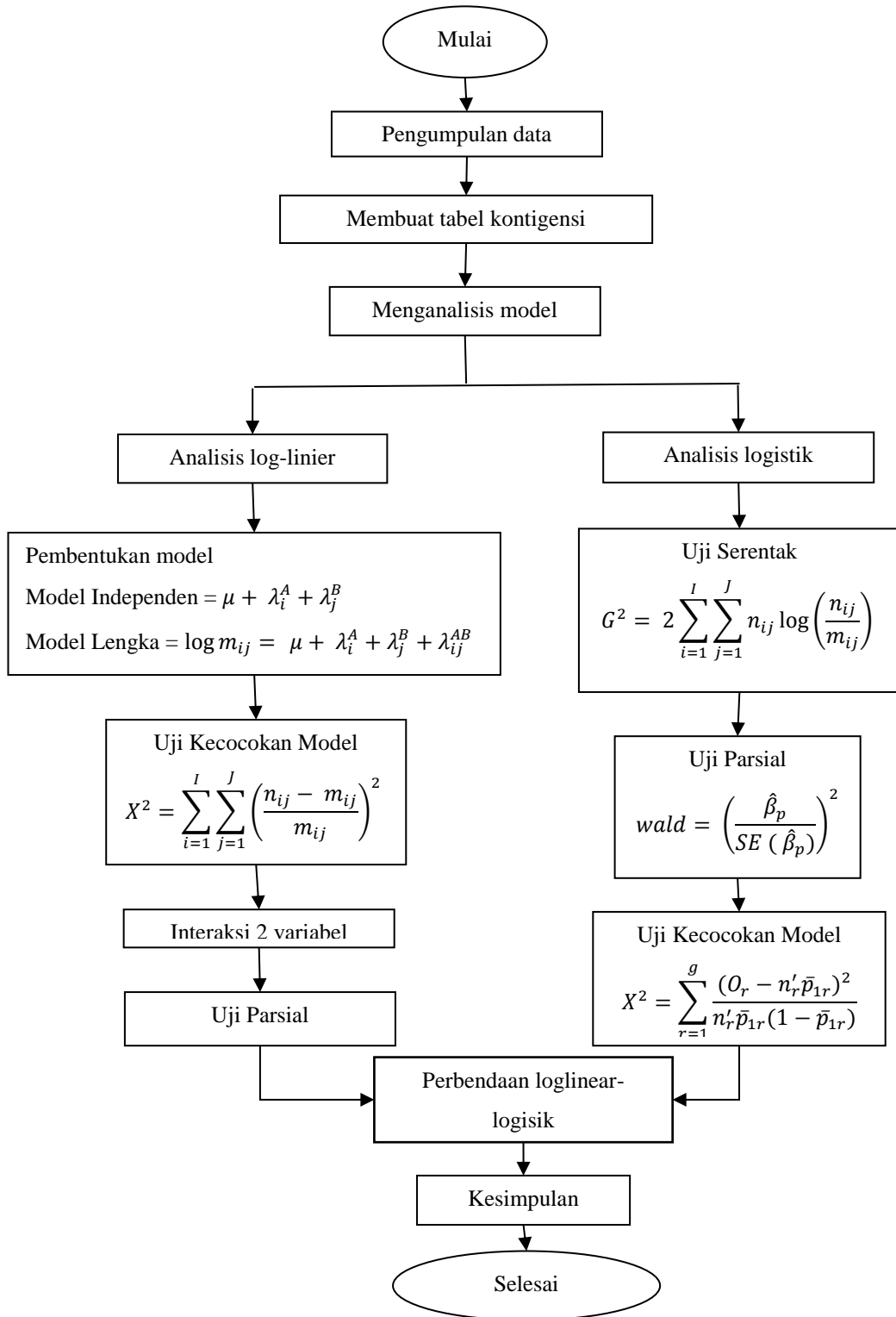
1. Membentuk model loglinear dari tabel dua dimensi untuk mencari model matematis secara pasti serta level mana yang cenderung menimbulkan adanya hubungan atau dependensi.
2. Melakukan uji goodness of fits dengan menggunakan uji chi-square Pearson dan ratio *Likelihood* untuk menguji hipotesis dari tiap model yang terbentuk.
3. Melakukan seleksi model terbaik
4. Membuat kesimpulan dan interpretasi dengan rasio odds dari model yang sudah diuji kecocokannya.

Metode analisis regresi logistik dengan langkah-langkah sebagai berikut.

1. Mengkategorikan variabel respon.
2. Membuat model regresi logistik secara parsial untuk setiap peubah bebas yang bertujuan untuk mengetahui ada tidaknya pengaruh dari variabel bebas terhadap variabel respon dengan menggunakan uji *rasio Likelihood*.
3. Meregresikan semua variabel bebas yang signifikan untuk mendapatkan model regresi logistik.

4. Melakukan pengujian hipotesis terhadap parameter regresi logistik yaitu pengujian secara serentak maupun parsial.
5. Membuat kesimpulan dan interpretasi dengan *rasio odds* dari model yang sudah diuji kecocokannya.

Diagram alir penelitian sebagaimana pada gambar (4.1)



**Gambar 4.1.** Diagram Alir Penelitian

## BAB V

### HASIL DAN PEMBAHASAN

#### 5.1 Analisis Model Loglinear pada tabel kontigensi dua arah

Loglinear digunakan untuk melihat sel-sel mana saja yang nantinya mempengaruhi dan menyebabkan hubungan antar variabelnya yang bersifat dependent. Selain itu analisis ini menggunakan data yang bersifat kualitatif (skala nominal atau ordinal).

Menurut (Agresti A. , 1990), data bivariat (A,B) dengan A dan B merupakan variabel kategorik, yang masing-masing mempunyai  $i$  baris dan  $j$  kolom. Misalnya ( $m_{ij}$ ) merupakan frekuensi ekspektasi dari sel (i,j) dari tabel silang berdimensi dua, maka model Loglinear mempunyai variabel respons yaitu  $\log(m_{ij})$ .

##### 5.1.1 Model Independent

Untuk model Loglinear independent yang mempunyai variabel respons yaitu  $\log(m_{ij})$  didapatkan model sebagai berikut (Nugraha, 2017):

$\log m_{ij} = \mu + \lambda_i^A + \lambda_j^B$  untuk  $i = 1 \dots I, j = 1 \dots J$ , atau

$$m_{ij} = \exp(\mu + \lambda_i^A + \lambda_j^B) \quad (5.1)$$

Keterangan:

$m_{ij}$  = frekuensi harapan

$\mu$  = parameter rata-rata

$\lambda_i^A$  = parameter pengaruh tingkat ke-i faktor pertama A

$\lambda_j^B$  = parameter pengaruh tingkat ke-i faktor kedua B

(Nugraha, 2019) Berdasarkan persamaan (5.1) maka didapatkan parameternya sebagai berikut :

a)  $\sum_{i=1}^b \log(m_{ij}) = \sum_{i=1}^b \log(n_{i+}) + b \log(n_{+j}) - b \log(n)$

$$\frac{1}{b} \sum_{i=1}^b \log(m_{ij}) = \frac{1}{b} \sum_{i=1}^b \log(n_{i+}) + \log(n_{+j}) - \log(n)$$

$$b) \sum_{j=1}^k \log(m_{ij}) = k \log(n_{i+}) + \sum_{j=1}^k \log(n_{+j}) - k \log(n)$$

$$\frac{1}{k} \sum_{j=1}^k \log(m_{ij}) = \log(n_{i+}) + \frac{1}{k} \sum_{j=1}^k \log(n_{+j}) - \log(n)$$

$$c) \sum_{i=1}^b \sum_{j=1}^k \log(m_{ij}) = \sum_{i=1}^b \sum_{j=1}^k \log(n_{i+}) + \sum_{i=1}^b \sum_{j=1}^k \log(n_{+j}) - \sum_{i=1}^b \sum_{j=1}^k \log(n)$$

$$\frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(m_{ij}) = \frac{1}{b} \sum_{i=1}^b \log(n_{i+}) + \frac{1}{k} \sum_{j=1}^k \log(n_{+j}) - \log(n)$$

Berdasarkan persamaan (a), (b), (c)

$$\log(\pi_{+j}) = \log\left(\frac{n_{+j}}{n}\right) = \log(n_{+j}) - \log(n)$$

$$\log(\pi_{i+}) = \log\left(\frac{n_{i+}}{n}\right) = \log(n_{i+}) - \log(n)$$

Maka didapatkan nilai persamaan estimasi sebagai berikut:

$$a) \mu = \log(n) + \frac{1}{b} \sum_{i=1}^b \log(\pi_{i+}) + \frac{1}{k} \sum_{j=1}^k \log(\pi_{+j})$$

$$\mu = \log(n) + \frac{1}{b} \sum_{i=1}^b \log(n_{i+}) - \log(n) + \frac{1}{k} \sum_{j=1}^k \log(n_{+j}) - \log(n)$$

$$\mu = \frac{1}{b} \sum_{i=1}^b \log(n_{i+}) + \frac{1}{k} \sum_{j=1}^k \log(n_{+j}) - \log(n)$$

$$b) \lambda_i^A = \log(\pi_{i+}) - \frac{1}{b} \sum_{i=1}^b \log(\pi_{i+})$$

$$\lambda_i^A = \log(n_{i+}) - \log(n) - \frac{1}{b} \sum_{i=1}^b \log(n_{i+}) + \log(n)$$

$$\lambda_i^A = \log(n_{i+}) - \frac{1}{b} \sum_{i=1}^b \log(n_{i+})$$

$$c) \lambda_i^B = \log(\pi_{+j}) - \frac{1}{k} \sum_{j=1}^k \log(\pi_{+j})$$

$$\lambda_i^B = \log(n_{+j}) - \log(n) - \frac{1}{k} \sum_{j=1}^k \log(n_{+j}) + \log(n)$$

$$\lambda_i^B = \log(n_{+j}) - \frac{1}{k} \sum_{j=1}^k \log(n_{+j})$$



### 5.1.2 Model Lengkap

Dimisalkan bahwa faktor A dan B bersifat dependent, maka model loglinear untuk model dua variabel sebagai berikut:

$$\log m_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \text{ untuk } i = 1, \dots, I, j = 1, \dots, J \quad (5.2)$$

(Nugraha, 2019) Berdasarkan persamaan (5.2) maka didapatkan parameternya sebagai berikut:

- a)  $\mu = \frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(m_{ij})$
- b)  $\lambda_i^A = \frac{1}{k} \sum_{j=1}^k \log(m_{ij}) - \frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(m_{ij})$
- c)  $\lambda_j^B = \frac{1}{b} \sum_{i=1}^b \log(m_{ij}) - \frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(m_{ij})$
- d)  $\lambda_{ij}^{AB} = \log(m_{ij}) - \frac{1}{k} \sum_{j=1}^k \log(m_{ij}) - \frac{1}{b} \sum_{i=1}^b \log(m_{ij}) - \frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(m_{ij})$

Maka dapat disusun model multinomial untuk masing-masing frekuensi sel. Dengan probabilitas (i,j) sebagai berikut:

$$\begin{aligned} \pi_{ij} &= \frac{m_{ij}}{\sum_{i=1}^b \sum_{j=1}^k m_{ij}} \\ &= \frac{\exp(\mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB})}{\sum_{i=1}^b \sum_{j=1}^k \exp(\mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB})} \end{aligned} \quad (5.3)$$

### 5.1.3 Persamaan Likelihood

Maximum *Likelihood* Estimator (MLE) digunakan untuk penaksiran paramater-parameter. Dimisalkan  $n_{i+}$  dan  $n_{+j}$  yang merupakan total frekuensi untuk baris ke-i dan kolom ke-j dan nilai n merupakan total observasi, maka penaksir untuk parameternya sebagai berikut: (Nugraha, 2019)

**Tabel 5.1.** Tabel persamaan *Likelihood*

Parameter	Model Independen	Model Lengkap
$\mu$	$\frac{1}{b} \sum_{i=1}^b \log(\hat{n}_{i+})$ $+ \frac{1}{k} \log(\hat{n}_{+j}) - \log(n)$	$\frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(\hat{m}_{ij})$

$\lambda_i^A$	$\log(\hat{n}_{i+})$ $-\frac{1}{b} \sum_{i=1}^b \log(\hat{n}_{i+})$	$\frac{1}{k} \sum_{j=1}^k \log(\hat{m}_{ij})$ $-\frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(\hat{m}_{ij})$
$\lambda_i^B$	$\log(\hat{n}_{+j})$ $-\frac{1}{k} \sum_{j=1}^k \log(\hat{n}_{+j})$	$\frac{1}{b} \sum_{i=1}^b \log(\hat{m}_{ij})$ $-\frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(\hat{m}_{ij})$
$\lambda_{ij}^{AB}$	-	$\log(\hat{m}_{ij})$ $-\frac{1}{k} \sum_{j=1}^k \log(\hat{m}_{ij})$ $-\frac{1}{b} \sum_{i=1}^b \log(\hat{m}_{ij})$ $-\frac{1}{bk} \sum_{i=1}^b \sum_{j=1}^k \log(\hat{m}_{ij})$

#### 5.1.4 Uji Kecocokan Model

Setelah mendapatkan estimasi harapan maka perlu membandingkan frekuensi dari hasil pengamatan dengan estimasi harapan untuk mengetahui apakah model loglinear tersebut cocok atau tidak. (Suryanto, 1988)

Berikut hipotesisnya:

$H_0$  = Model Loglinear layak digunakan

$H_1$  = Model Loglinear tidak layak digunakan

Dalam pengujian hipotesis estimasi frekuensi harapan layak atau tidak digunakan statistik uji *Chi-square*, sebagai berikut:

$$X^2 = \sum_{i=1}^I \sum_{j=1}^J \left( \frac{n_{ij} - m_{ij}}{m_{ij}} \right)^2 \quad (5.4)$$

Apabila  $\chi_{hitung}^2 \leq \chi_{tabel}^2$  atau gagal menolak  $H_0$ , dengan menggunakan taraf signifikansi yaitu  $\alpha = 0,05$  maka model Loglinear yang digunakan sesuai atau layak.

Selain statistik *Chi-square*, dapat juga menggunakan statistik rasio *Likelihood*, sebagai berikut:

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log \left( \frac{n_{ij}}{m_{ij}} \right) \quad (5.5)$$

Untuk menentukan derajat bebas untuk uji *Goodness of fit* didapatkan dari selisih antara jumlah sel yang bebas dengan model yang ditentukan

### 5.1.5 Partisi *Chi-square* untuk Membandingkan Model

Permisalan didapatkan dua model parametrik yaitu  $m_1$  dan  $m_2$  dengan  $m_2$  merupakan kasus khusus dari  $m_1$ , dikarenakan dalam hal ini  $m_2$  lebih sederhana dibandingkan  $m_1$ , maka diperoleh model  $m_2$  bersusun dengan  $m_1$ , selanjutnya untuk  $v_1$  dan  $v_2$  adalah derajat bebas sesatan dan  $v_1$  nilainya lebih kecil dibandingkan  $v_2$ . Maka diperoleh:

$$G^2(m_1) \leq G^2(m_2) \quad (5.6)$$

Pada persamaan (5.6) dapat kita lihat bahwa secara teoritis,  $G^2(m_1)$  tidak melampaui  $G^2(m_2)$ , diasumsikan bahwa model  $m_1$  sudah ditentukan. Maka dalam pendekatan rasio *Likelihood* untuk menguji apakah  $m_2$  dapat dihitung dengan uji statistic yaitu:

$$G^2(m_2) = G^2(m_1) + G^2(m_1 + m_2) \quad (5.7)$$

Untuk  $G^2(m_1)$  mendekati distribusi *Chi-square* dengan derajat bebas  $v_1$ ,  $G^2(m_2)$  mendekati distribusi *Chi-square* dengan derajat bebas  $v_1$ . Oleh sebab itu, diperoleh  $G^2(m_2|m_1)$  mendekati distribusi *Chi-square* dengan derajat bebas  $v_2 - v_1$ .

## 5.2 Penerapan Model Loglinear

Penerapan Model Loglinear pada tabel kontigensi 2 arah dengan menggunakan data kasus Covid-19 tahun 2020 menurut negara dan status keadaan. Berdasarkan dari data kasus Covid-19 diperoleh tabel yang memuat data 23007 jiwa orang yang diklasifikasikan berdasarkan tiga kategori yaitu sembuh, meninggal dan dalam perawatan dari dua negara yang berbeda, didapatkan data sebagai berikut:

**Tabel 5.2.** Tabel data kasus covid-19

Negara	Status		
	Sembuh	Meninggal	Perawatan
Indonesia	1608	743	7032
Singapura	1060	12	12552

### 5.2.1 Menentukan Model Loglinear dan Loglinear penuh

Berdasarkan studi kasus pada tabel 5.2 maka diperoleh model loglinear dan loglinear penuh sebagai berikut:

Model Loglinear:

$$\ln m_{ij} = \mu + \lambda_i^A + \lambda_j^B \quad (i = 1,2 ; j = 1,2)$$

Sedangkan untuk model loglinear penuh sebagai berikut:

$$\ln m_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \quad (i = 1,2 ; j = 1,2)$$

Keterangan:

$m_{ij}$  = Frekuensi harapan dalam sel ke-ij

$\mu$  = Rataan umum

$\lambda_i^A$  = Parameter pengaruh utama kategori ke-i dari variabel Status

$\lambda_j^B$  = Parameter pengaruh utama kategori ke-j dari variabel Negara

$\lambda_{ij}^{AB}$  = Parameter pengaruh interaksi antara Status dengan Negara

Dengan syarat sebagai berikut:

$$\sum_{i=1}^2 \lambda_i^A = 0, \sum_{j=1}^2 \lambda_j^B = 0, \sum_{i=1}^2 \lambda_{ij}^{AB} = 0, \sum_{j=1}^2 \lambda_{ij}^{AB} = 0$$

### 5.2.2 Pendugaan Nilai Parameter

Dalam pendugaan nilai parameter didapatkan dengan cara mensubstitusikan nilai observasi pada tabel ke dalam persamaan sebelumnya. Sehingga di dapatkan nilai pendugaan parameter sebagai berikut:

$$\mu = \frac{\ln(1608) + \ln(743) + \ln(7032) + \ln(1060) + \ln(12) + \ln(12552)}{6}$$

$$= 6,957$$

$$\lambda_{A(1)} = \frac{\ln(1608) + \ln(743) + \ln(7032)}{3} - \mu = 0,6605$$

$$\lambda_{A(2)} = \frac{\ln(1060) + \ln(12) + \ln(12552)}{3} - \mu = -0,6605$$

$$\lambda_{A(3)} = \frac{\ln(7032) + \ln(12552)}{2} - \mu = 2,1912$$

$$\lambda_{B(1)} = \frac{\ln(1060) + \ln(1060)}{2} - \mu = 0,2177$$

$$\lambda_{B(2)} = \frac{\ln(743) + \ln(12)}{2} - \mu = -2,4089$$

Berdasarkan hasil estimasi diatas menunjukkan bahwa semua nilai penduga parameter merupakan pengaruh utama. Untuk parameter  $\lambda_{A(j)}$  yang mempunyai nilai penduga terbesar adalah  $\lambda_{A(3)}$ , karena mempunyai nilai total marjinal paling besar diantara variabel A. Sedangkan untuk parameter  $\lambda_{B(j)}$  yang mempunyai nilai penduga terbesar adalah  $\lambda_{B(1)}$ , Selanjutnya untuk penduga pengaruh interaksi sebagai berikut:

$$\lambda_{AB(11)} = \ln(1608) - \lambda_{A(1)} - \lambda_{B(1)} - \mu = -0,452$$

$$\lambda_{AB(12)} = \ln(1060) - \lambda_{A(1)} - \lambda_{B(2)} - \mu = 1,758$$

$$\lambda_{AB(21)} = \ln(743) - \lambda_{A(2)} - \lambda_{B(1)} - \mu = 0,097$$

$$\lambda_{AB(22)} = \ln(12) - \lambda_{A(2)} - \lambda_{B(2)} - \mu = -1,402$$

$$\lambda_{AB(31)} = \ln(7032) - \lambda_{A(3)} - \lambda_{B(1)} - \mu = -0,507$$

$$\lambda_{AB(32)} = \ln(12552) - \lambda_{A(3)} - \lambda_{B(2)} - \mu = 2,699$$

### 5.2.3 Uji Kecocokan Model

Untuk menguji model manakah yang sesuai, maka digunakan statistik uji chi-square dan rasio *Likelihood*. Selanjutnya untuk mencari nilai  $\chi^2$  dan  $G^2$  harus mencari nilai harapan untuk setiap sel terlebih dahulu, dengan persamaan sebagai berikut:

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

Maka didapatkan nilai harapan sebagai berikut:

**Tabel 5.3.** Tabel estimasi nilai harapan

Negara	Status		
	Sembuh	Meninggal	Perawatan
Indonesia	1088,10	307,91	7986,99
Singapura	1579,90	447,09	11597,01

Statistik uji *Chi-square*, sebagai berikut:

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \left( \frac{n_{ij} - e_{ij}}{e_{ij}} \right)^2 \quad (5.8)$$

$$\begin{aligned} \chi^2 &= \frac{(n_{11} - e_{11})^2}{e_{11}} + \frac{(n_{12} - e_{12})^2}{e_{12}} + \dots + \frac{(n_{23} - e_{23})^2}{e_{23}} \\ &= \frac{(1608 - 1088,10)^2}{1088,10} + \frac{(743 - 307,91)^2}{307,91} + \dots + \frac{(19584 - 11597,01)^2}{11597,01} \\ &= 1650,52 \end{aligned}$$

Berikut merupakan hasil untuk Statistik rasio *Likelihood*:

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log \left( \frac{n_{ij}}{e_{ij}} \right) \quad (5.9)$$

$$\begin{aligned} G^2 &= 2 \left[ \left[ n_{11} \ln \left( \frac{n_{11}}{e_{11}} \right) \right] + \left[ n_{12} \ln \left( \frac{n_{12}}{e_{12}} \right) \right] + \dots + \left[ n_{23} \ln \left( \frac{n_{23}}{e_{23}} \right) \right] \right] \\ &= 2 \left[ \left[ 1608 \ln \left( \frac{1608}{1088,10} \right) \right] + \left[ 743 \ln \left( \frac{743}{307,91} \right) \right] + \dots \right. \\ &\quad \left. + \left[ 19584 \ln \left( \frac{19584}{11597,01} \right) \right] \right] \\ &= 1827,72 \end{aligned}$$

a) Hipotesis

$$H_0 : \text{Model loglinear } (\pi_{ij} = \mu + \lambda_i^A + \lambda_j^B)$$

$$H_1 : \text{Model loglinear penuh } (\pi_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB})$$

b) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

c) Daerah Kritis

$H_0$  ditolak jika  $\chi^2 > \chi_{tabel}^2$  dan  $G^2 > \chi_{tabel}^2$

d) Statistik Uji

$$\chi^2 = 1650,52$$

$$G^2 = 1827,72$$

$$\chi_{tabel}^2 = 5,991$$

e) Keputusan

Tolak  $H_0$ , karena nilai  $1650,52 > 5,991$  dan  $1827,72 > 5,991$

f) Kesimpulan

Dengan menggunakan tingkat kepercayaan 95% data yang ada menolak  $H_0$ , yang artinya bahwa model lengkap ( $\pi_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$ ) adalah model terbaik. Selanjutnya, maka perlu diuji interaksi 2 variabel.

#### 5.2.4 Interaksi dua variabel

Setelah mengetahui bahwa model terbaik yang digunakan adalah model lengkap, maka selanjutnya peneliti akan melakukan pengujian interaksi dari dua variabel. Analisis ini dilakukan untuk mengetahui pengaruh interaksi dari orde ke-k. Berikut analisisnya:

**Tabel 5.4.** Tabel uji interaksi dua variabel

	K	Df	Likelihood Ratio	Pearson
K-way and Higher Order	1	5	30202,817	32091,519
	2	2	1827,875	1650,698
K-way Effects	1	3	28374,941	20440,821
	2	2	1827,875	1650,698

#### i. Untuk k faktor yang lebih tinggi sama dengan nol

a) Hipotesis

$H_0 : \lambda^A = 0; \lambda^B = 0;$  (Tidak ada pengaruh interaksi dari orde ke -k dan yang lebih tinggi)

$H_1 : \text{Minimal terdapat satu diantara } \lambda^A \text{ atau } \lambda^B \text{ yang } \neq 0$  (Ada pengaruh interaksi dari orde ke-k dan yang lebih tinggi)

b) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

c) Daerah Kritis

$$H_0 \text{ ditolak jika } \chi^2 > \chi_{tabel}^2$$

d) Statistik Uji

Menggunakan metode *Likelihood Ratio* dilihat dari K-way and Higher Order diperoleh:

$$X_{hitung}^2 = 30202,817 \text{ ( k = 1)}$$

$$X_{hitung}^2 = 1827,875 \text{ ( k=2 )}$$

e) Keputusan

Tolak  $H_0$  karena  $X_{hitung}^2$  (30202,817)  $>$   $X_{tabel}^2$  (11,070) dan  $X_{hitung}^2$ (1827,875)  $>$   $X_{tabel}^2$ (5,991)

f) Kesimpulan

Dengan tingkat kepercayaan 95% maka keputusan untuk k=1 dan k=2 menolak  $H_0$  yang artinya bahwa secara signifikan menjelaskan hubungan antar peubah untuk semua tingkat interaksi ada dalam model dan minimal interaksi 2 faktor harus terdapat dalam model.

## ii. Untuk k faktor sama dengan nol

a) Hipotesis

$$H_0 : \lambda^{AB} = 0; \text{ (Tidak ada pengaruh interaksi dari orde ke-k)}$$

$$H_1: \lambda^{AB} \neq 0 \text{ (Ada pengaruh interaksi dari orde ke-k)}$$

b) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

c) Daerah kritis

$$H_0 \text{ ditolak jika } \chi^2 > \chi_{tabel}^2$$

d) Statistik Uji

Menggunakan metode *Likelihood Ratio* dilihat dari K-way *Effects* diperoleh:

$$X_{hitung}^2 = 28374,941 \text{ ( k = 1)}$$

$$X_{hitung}^2 = 1827,875 \text{ ( k=2 )}$$



e) Keputusan

Tolak  $H_0$  karena  $X_{hitung}^2 (28374,941) > X_{tabel}^2 (7,815)$  dan  $X_{hitung}^2(1827,875) > X_{tabel}^2(5,991)$

f) Kesimpulan

Dengan tingkat kepercayaan 95% maka keputusan untuk  $k=1$  dan  $k=2$  menolak  $H_0$  yang artinya bahwa secara signifikan menjelaskan hubungan antar peubah untuk semua tingkat interaksi ada dalam model dan minimal interaksi 2 faktor harus terdapat dalam model.

### 5.2.5 Uji Parsial

Selanjutnya adalah menguji kebebasan secara parsial yang nantinya akan menunjukkan interaksi-interaksi yang ada dalam model.

**Tabel 5.5.** Tabel uji parsial

Effect	Df	Partial Chi-Square
Negara	1	786,593
Status	2	27588,349

#### i. Faktor Negara

a) Hipotesis

$H_0$  : Tidak ada pengaruh efek faktor negara

$H_1$  : Ada pengaruh efek faktor negara

b) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

c) Daerah Kritis

$H_0$  ditolak jika  $X_{hitung}^2 > X_{tabel}^2$

d) Statistik Uji

Menggunakan metode *Likelihood Ratio* diperoleh:  $X_{hitung}^2 = 786,593$

e) Keputusan

Tolak  $H_0$  karena  $X_{hitung}^2(786,593) > X_{tabel}^2(3,841)$

f) Kesimpulan

Dengan tingkat kepercayaan 95% maka keputusan yang ada menolak  $H_0$  yang artinya terdapat pengaruh faktor negara.

## ii. Faktor Status

a) Hipotesis

$H_0$  : Tidak ada pengaruh efek faktor status

$H_1$ : Ada pengaruh efek faktor status

b) Tingkat Signifikansi

$\alpha = 5\% = 0.05$

c) Daerah kritis

$H_0$  ditolak jika  $X_{hitung}^2 > X_{tabel}^2$

d) Statistik Uji

Menggunakan metode *Likelihood Ratio* diperoleh:  $X_{hitung}^2 = 27588,349$

e) Keputusan

Tolak  $H_0$  karena  $X_{hitung}^2(27588,349) > X_{tabel}^2(5,991)$

f) Kesimpulan

Dengan tingkat kepercayaan 95% maka keputusan yang ada menolak  $H_0$  yang artinya terdapat pengaruh efek factor status.

Berdasarkan analisis diatas maka dihasilkan dua efek utama yaitu faktor negara dan status. Apabila diurutkan berdasarkan tingkat signifikansinya.

## 5.3 Regresi Logistik

Regresi logistik digunakan untuk mencari hubungan antar variabel respon yang memiliki skala nominal atau ordinal dengan dua kategori atau lebih dari dua kategori dengan satu atau lebih variabel prediktor dan variable respon.

### 5.3.1 Regresi Logistik Multinomial

Regresi logistik multinomial digunakan untuk variabel dependen yang mempunyai skala yang bersifat multinomial. Sedangkan metode yang digunakan dalam penelitian ini adalah regresi logistik dengan variabel respon berskala nominal dengan tiga kategori.

Model yang digunakan pada regresi logistik multinomial adalah (Agresti A. , 1990)

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)} \quad (0.10)$$

Apabila variabel respon memiliki tiga kategori maka akan membentuk dua persamaan logit, dan masing-masing persamaan akan membentuk regresi logistik multinomial dengan membandingkan suatu kelompok kategori. Berikut fungsi logit :

$$\begin{aligned} g_1(x) &= \ln \left[ \frac{P(Y = 1)|x}{P(Y = 3)|x} \right] \\ &= \beta_{10} + \beta_{11}X_1 + \beta_{12}X_2 + \dots + \beta_{1p}X_p \\ &= x' \beta_1 \end{aligned}$$

$$\begin{aligned} g_2(x) &= \ln \left[ \frac{P(Y = 2)|x}{P(Y = 3)|x} \right] \\ &= \beta_{20} + \beta_{21}X_1 + \beta_{22}X_2 + \dots + \beta_{2p}X_p \\ &= x' \beta_2 \end{aligned}$$

Berdasarkan kedua fungsi logit didapatkan peluang untuk masing-masing kategori respon sebagai berikut :

$$\pi_1(x) = \frac{\exp g_1(x)}{1 + \exp g_1(x) + \exp g_2(x)} \quad (0.11)$$

$$\pi_2(x) = \frac{\exp g_2(x)}{1 + \exp g_1(x) + \exp g_2(x)} \quad (0.12)$$

$$\pi_3(x) = \frac{1}{1 + \exp g_1(x) + \exp g_2(x)} \quad (0.13)$$

Dalam pengujian signifikansi  $\beta$  dari model yang telah diperoleh, maka dilakukan uji serentak dan uji parsial

### 5.3.2 Uji Serentak

Uji serentak menggunakan Statistik uji  $G^2$  atau *Likelihood* ratio test yang dinyatakan sebagai berikut:

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log \left( \frac{n_{ij}}{m_{ij}} \right) \quad (0.14)$$

Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \beta_p = 0$$

$H_1$  : Paling tidak ada satu parameter  $\beta_p \neq 0$

Menurut Hosmer dan Lemeshow (1989), statistik uji  $G^2$  mengikuti distribusi chi-square, sehingga untuk memperoleh keputusan dilakukan perbandingan dengan titik kritis  $\chi^2_{(\alpha, db)}$  di mana derajat bebasnya adalah  $p$  atau banyaknya variabel prediktor. Kriteria penolakan (tolak  $H_0$ ) jika nilai  $G^2 > \chi^2_{(\alpha, db)}$

### 5.3.3 Uji Parsial

Pengujian signifikansi parameter menggunakan uji Wald. Uji Wald digunakan untuk menguji signifikansi model secara terpisah atau digunakan untuk mengetahui variabel prediktor yang berpengaruh terhadap model (Hosmer, 1989). Hipotesis yang digunakan adalah sebagai berikut:

$H_0$  :  $\beta_p = 0$ , dengan  $p = 1, 2, \dots, p$  (tidak ada pengaruh variabel bebas ke- $p$  terhadap variabel terikat)

$H_1$ :  $\beta_p \neq 0$ , dengan  $p=1, 2, \dots, p$  (ada pengaruh variabel bebas ke- $p$  terhadap variabel terikat)

Perhitungan statistik untuk uji Wald ( $W$ ) :

$$wald = \left( \frac{\hat{\beta}_p}{SE(\hat{\beta}_p)} \right)^2 \quad (0.15)$$

Rasio yang dihasilkan untuk uji statistik dibawah hipotesis  $H_0$  , maka akan mengikuti sebaran normal baku, sehingga keputusan didapatkan dari perbandingan dengan distribusi normal baku ( $Z$ ). Kreteria penolakan adalah tolak  $H_0$  jika nilai  $W > Z_{\alpha/2}$  atau  $p\text{-value} < \alpha$ .

### 5.3.4 Uji Kesesuaian Model

Hipotesis

$H_0$  : Model sesuai (Tidak ada perbedaan yang nyata antara hasil observasi dengan kemungkinan hasil prediksi model)

$H_1$  : Model tidak sesuai (ada perbedaan yang nyata antara hasil observasi dengan kemungkinan hasil prediksi model)

Uji kesesuaian model dengan menggunakan statistik uji Chi-square adalah sebagai berikut: (Hosmer, 1989)

$$\chi^2 = \sum_{r=1}^g \frac{(O_r - n'_r \bar{p}_{1r})^2}{n'_r \bar{p}_{1r} (1 - \bar{p}_{1r})} \quad (0.16)$$

Pengambilan keputusan didasarkan pada tolak  $H_0$  jika  $\chi^2_{(hitung)} \geq \chi^2_{(\alpha, db)}$

Dalam menginterpretasi model logistik dapat menggunakan nilai *odds ratio*, berikut persamaan nya:

$$OR_j(a, b) = \frac{P(Y = j|x = a) / P(Y = 1|x = a)}{P(Y = j|x = b) / P(Y = 1|x = b)} \quad (0.17)$$

Terdapat hubungan nilai odds ratio dengan parameter model ( $\beta$ ) yaitu:  $OR_j(a, b) = \exp(\hat{\beta})$

## 5.4 Penerapan Model Logistik Multinomial

### 5.4.1 Pengujian Serentak

Statistik pengujian serentak (Uji G) bertujuan untuk mengetahui pengaruh variabel bebas secara serentak terhadap variabel terikat.

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log \left( \frac{n_{ij}}{m_{ij}} \right)$$

$$G^2 = 2 \left[ \left[ n_{11} \ln \left( \frac{n_{11}}{e_{11}} \right) \right] + \left[ n_{12} \ln \left( \frac{n_{12}}{e_{12}} \right) \right] + \dots + \left[ n_{23} \ln \left( \frac{n_{23}}{e_{23}} \right) \right] \right]$$

$$= 2 \left[ \left[ 1608 \ln \left( \frac{1608}{1088,10} \right) \right] + \left[ 743 \ln \left( \frac{743}{307,91} \right) \right] + \dots \right.$$

$$\left. + \left[ 19584 \ln \left( \frac{19584}{11597,01} \right) \right] \right]$$

$$= 1827,72$$

**Tabel 5.6.** Likelihood Ratio Test

Likelihood Ratio	1827,72
------------------	---------

a) Hipotesis

$H_0 : \beta_1 = \beta_2 = \beta_p = 0$  (Tidak ada pengaruh variabel prediktor terhadap model)

$H_1$ : Paling tidak ada satu  $\beta_p$  yang tidak sama dengan 0

b) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

c) Daerah kritis

$$H_0 \text{ ditolak jika } X_{hitung}^2 > X_{tabel}^2$$

d) Statistik Uji

$$X_{hitung}^2 = 1827.72$$

e) Keputusan

$$\text{Tolak } H_0 \text{ karena } X_{hitung}^2(1827,875) > X_{tabel}^2(5,991)$$

f) Kesimpulan

Dengan menggunakan tingkat kepercayaan 95% maka dapat disimpulkan bahwa minimal ada satu variabel *independent* yang signifikan mempengaruhi variabel *dependent*. atau dengan kata lain secara keseluruhan terdapat pengaruh antara Status (Y) dengan variabel independen Negara (X).

#### 5.4.2 Pengujian Parsial

Uji *Wald* digunakan untuk mengetahui pengaruh masing-masing variabel *independen* terhadap variabel *dependen*.

**Tabel 5.7.** Hasil Estimasi Parameter

Logit	Variabel Prediktor	Koefisien	SE	Wald	P-value	OR
1	Konstanta	-2,472	0,032	5971,138	0,000	
	Negara	0,996	0,042	555,225	0,000	2,708
2	Konstanta	-6,953	0,289	579,531	0,000	
	Negara	4,705	0,291	260,761	0,000	110,520

Pengujian hipotesis untuk hasil pada tabel **5.7** adalah sebagai berikut :

Hipotesis

$H_0 : \beta_p = 0$ , dengan  $p = 1, 2, \dots, p$  (tidak ada pengaruh variable bebas ke- $p$  terhadap variable terikat)

$H_1 : \beta_p \neq 0$ , dengan  $p = 1, 2, \dots, p$  (ada pengaruh variabel bebas ke- $p$  terhadap variable terikat)

Berdasarkan hasil estimasi variable pada tabel (5.7) dapat dilihat bahwa semua variabel yang ada menunjukkan bahwa variabel prediktor berpengaruh signifikan terhadap model. Karena nilai  $W_k > Z_{\alpha/2}$  atau nilai p-value  $< \alpha$ .

### 5.4.3 Pembentukan Model Logistik Multinomial

Setelah didapatkan kesimpulan bahwa variabel independen signifikan mempengaruhi variabel dependent. Selanjutnya peneliti akan membentuk model Regresi Logit dengan hasil estimasi berdasarkan hasil tabel (5.7) seperti berikut :

Model Logit 1

$$\begin{aligned} g_1(x) &= \ln \left[ \frac{P(Y = 1)x}{P(Y = 3)x} \right] \\ &= \beta_{10} + \beta_{11}X \\ &= -2,472 + 0,996 X \end{aligned}$$

Model Logit 2

$$\begin{aligned} g_2(x) &= \ln \left[ \frac{P(Y = 2)x}{P(Y = 3)x} \right] \\ &= \beta_{20} + \beta_{22}X \\ &= -6,953 + 4,705 X \end{aligned}$$

Pada Tabel (5.7) terdapat nilai *odds ratio*. Berikut adalah perhitungan manual untuk mencari nilai *odds ratio* tabel kontingensi 2 arah :

- i. Nilai *odds ratio* variabel Sembuh dan Meninggal

$$\frac{\mu_{11}\mu_{22}}{\mu_{21}\mu_{12}} = \frac{(1608)(12)}{(1060)(743)} = 0,0245$$

- ii. Nilai *odds ratio* variabel Sembuh dan Perawatan

$$\frac{\mu_{11}\mu_{23}}{\mu_{21}\mu_{13}} = \frac{(1608)(12552)}{(1060)(7032)} = 2,708$$

- iii. Nilai *odds ratio* variabel Meninggal dan Perawatan

$$\frac{\mu_{12}\mu_{23}}{\mu_{22}\mu_{13}} = \frac{(743)(12552)}{(12)(7032)} = 110,52$$

Interpretasi nilai *odds ratio* untuk masing-masing kategori variabel prediktor sesuai tabel (5.7). Dapat diketahui bahwa nilai odds ratio pada logit 1

untuk kategori status Sembuh adalah sebesar 2,708 yang berarti suatu Negara memiliki angka kejadian untuk seseorang dalam perawatan sebesar 2,708 kali lebih besar dibandingkan dengan seseorang yang berstatus dalam perawatan di suatu Negara. Selanjutnya pada Logit 2 yang menjelaskan kategori status Meninggal adalah sebesar 110,52 yang berarti suatu Negara memiliki angka kejadian untuk seseorang meninggal sebesar 110,52 kali lebih besar dibandingkan dengan seseorang yang berstatus dalam perawatan di suatu negara

#### 5.4.4 Uji kecocokan model

Uji kecocokan model merupakan uji statistik yang digunakan untuk menguji ada atau tidak ada perbedaan yang signifikan antara model yang diperoleh dengan data yang digunakan oleh peneliti.

Berikut statistik uji dengan menggunakan uji Chi-Square:

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \left( \frac{n_{ij} - e_{ij}}{e_{ij}} \right)^2$$

$$\chi^2 = \frac{(n_{11} - e_{11})^2}{e_{11}} + \frac{(n_{12} - e_{12})^2}{e_{12}} + \dots + \frac{(n_{23} - e_{23})^2}{e_{23}}$$

$$= \frac{(1608 - 1088,10)^2}{1088,10} + \frac{(743 - 307,91)^2}{307,91} + \dots + \frac{(19584 - 11597,01)^2}{11597,01}$$

$$= 1650,52$$

a) Hipotesis

$H_0$  : Tidak ada perbedaan antara hasil observasi dengan kemungkinan prediksi model.

$H_1$  : Ada perbedaan antara hasil observasi dengan kemungkinan model.

b) Tingkat Signifikansi

$$\alpha = 5\% = 0.05$$

c) Daerah kritis

$$H_0 \text{ ditolak jika } X_{hitung}^2 > X_{tabel}^2$$

d) Statistik Uji

$$X_{hitung}^2 = 1650,52$$



e) Keputusan

Tolak  $H_0$  karena  $X^2_{hitung}(1650,52) > X^2_{(db,\alpha)}(5,991)$

f) Kesimpulan

Dengan menggunakan tingkat kepercayaan 95% maka dapat disimpulkan bahwa ada perbedaan antara hasil observasi dengan kemungkinan model

## 5.5 Menentukan nilai Prediksi Frekuensi Harapan

Pada pembahasan kali peneliti akan menentukan nilai prediksi frekuensi harapan masing-masing sel dengan data observasi yang ada. Setelah itu peneliti bandingkan melalui persamaan tersebut lebih dekat persamaan loglinear atau logistik.

### 5.5.1 Loglinear

Untuk menentukan nilai prediksi frekuensi harapan masing-masing sel pada model loglinear peneliti akan membuat estimasi parameter loglinear terlebih dahulu dengan menggunakan tabel (5.8) berikut ini.

**Tabel 5.8.** Estimasi Parameter Loglinear

Variabel	Parameter	Estimasi	P-Value	Keputusan
Negara*Status	1	-0,446	0,000	Sig.
	2	1,389	0,000	Sig.
Negara	1	0,654	0,000	Sig.
Status	1	0,211	0,000	Sig.
	2	-2,395	0,000	Sig.

Berdasarkan tabel estimasi parameter untuk loglinear maka didapatkan persamaan sebagai berikut:

$$\log(\pi_{ij}) = \lambda + \lambda^A + \lambda_1^B - \lambda_2^B - \lambda_1^{AB} + \lambda_2^{AB}$$

Berikut merupakan rumus untuk mencari prediksi frekuensi

$$\pi_{ij} = \frac{\exp(\lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB})}{n}$$

$n$  merupakan jumlah pengamatan.

Selanjutnya peneliti akan membuat ilustrasi perhitungan tingkat status (sembuh, meninggal, dalam perawatan) dengan menggunakan negara. Berikut ilustrasi perhitungan status keadaan dikarenakan Covid-19.

Berikut adalah hasil perhitungan berdasarkan prediksi frekuensi:

$$\pi_{ij} = \frac{\exp((0,654) + 0,211 + (-2,395) + (-0,446) + 1,389)}{n}$$

$$\pi_{ij} = \frac{\exp(-0.587)}{1} = 0,587$$

Berdasarkan hasil nilai peluang pengamatan pada baris ke i dan kolom ke j, maka disimpulkan bahwa untuk peluang korban jiwa pada negara dengan berstatus sembuh, meninggal dan dalam perawatan sebesar 0,587%.

### 5.5.2 Logistik

Untuk menentukan nilai prediksi frekuensi harapan masing-masing sel pada model logistik peneliti akan membuat estimasi parameter logistik terlebih dahulu dengan menggunakan tabel (5.9) berikut ini.

**Tabel 5.9.** Estimasi Parameter Logistik

Logit	Variabel Prediktor	Koefisen	SE	Wald	P-value	OR
1	Konstanta	-2,472	0,032	5971,138	0,000	
	Negara	0,996	0,042	555,225	0,000	2,708
2	Konstanta	-6,953	0,289	579,531	0,000	
	Negara	4,705	0,291	260,761	0,000	110,520

Seperti yang sudah peneliti jelaskan pada sub bab sebelumnya yaitu mengenai pembentukan dalam model regresi logistik multinomial didapatkan model signifikan pada model logit 1 dan model logit 2. Selanjutnya peneliti akan membuat ilustrasi perhitungan tingkat status (sembuh, meninggal, dalam perawatan). Berikut ilustrasi perhitungan status keadaan dikarenakan Covid-19.

Berikut untuk menentukan peluang dari masing-masing sel.

1. Untuk  $x=0$  (Negara Indonesia)

Model Logit 1 (Sembuh)

$$\begin{aligned}
g_1(x) &= \ln \left[ \frac{P(Y = 1)x}{P(Y = 3)x} \right] \\
&= \beta_{10} + \beta_{11}X \\
&= -2,472 + 0,996 (0) \\
&= -2,472
\end{aligned}$$

Model Logit 2 (Meninggal)

$$\begin{aligned}
g_2(x) &= \ln \left[ \frac{P(Y = 2)x}{P(Y = 3)x} \right] \\
&= \beta_{20} + \beta_{22}X \\
&= -6,953 + 4,705 (0) \\
&= -6,953
\end{aligned}$$

Sehingga didapatkan nilai peluang masing-masing kategori sebagai berikut:

$$\begin{aligned}
\pi_1(x) &= \frac{\exp g_1(x)}{1 + \exp g_1(x) + \exp g_2(x)} \\
&= \frac{\exp(-2,472)}{1 + \exp(-2,472) + \exp(-6,953)} = 0,078
\end{aligned}$$

$$\begin{aligned}
\pi_2(x) &= \frac{\exp g_2(x)}{1 + \exp g_1(x) + \exp g_2(x)} \\
&= \frac{\exp(-2,473)}{1 + \exp(-2,473) + \exp(-6,953)} = 0,001
\end{aligned}$$

$$\begin{aligned}
\pi_3(x) &= \frac{1}{1 + \exp g_1(x) + \exp g_2(x)} \\
&= \frac{1}{1 + \exp(-2,473) + \exp(-6,953)} = 0,921
\end{aligned}$$

Berdasarkan nilai peluang masing-masing model logit, maka disimpulkan bahwa untuk peluang korban jiwa pada negara Indonesia dengan berstatus sembuh sebesar 0,078% , untuk peluang korban jiwa pada negara Indonesia dengan berstatus dalam meninggal sebesar 0,001% dan peluang korban jiwa pada negara Indonesia dengan berstatus dalam perawatan sebesar 0,921%

2. Untuk x=1 (Negara Singapura)

Model Logit 1 (Sembuh)

$$\begin{aligned}g_1(x) &= \ln \left[ \frac{P(Y = 1)x}{P(Y = 3)x} \right] \\ &= \beta_{10} + \beta_{11}X \\ &= -2,472 + 0,996 x \\ &= -1,476\end{aligned}$$

Model Logit 2 (Meninggal)

$$\begin{aligned}g_2(x) &= \ln \left[ \frac{P(Y = 2)x}{P(Y = 3)x} \right] \\ &= \beta_{20} + \beta_{22}X \\ &= -6,953 + 4,705 x \\ &= -2,248\end{aligned}$$

Sehingga didapatkan nilai peluang masing-masing kategori sebagai berikut:

$$\begin{aligned}\pi_1(x) &= \frac{\exp g_1(x)}{1 + \exp g_1(x) + \exp g_2(x)} \\ &= \frac{\exp(-1,476)}{1 + \exp(-1,476) + \exp(-2,248)} = 0,171\end{aligned}$$

$$\begin{aligned}\pi_2(x) &= \frac{\exp g_2(x)}{1 + \exp g_1(x) + \exp g_2(x)} \\ &= \frac{\exp(-2,248)}{1 + \exp(-1,476) + \exp(-2,248)} = 0,079\end{aligned}$$

$$\begin{aligned}\pi_3(x) &= \frac{1}{1 + \exp g_1(x) + \exp g_2(x)} \\ &= \frac{1}{1 + \exp(-1,476) + \exp(-2,248)} = 0,750\end{aligned}$$

Berdasarkan nilai peluang masing-masing model logit, maka disimpulkan bahwa untuk peluang korban jiwa pada negara Singapura dengan berstatus sembuh sebesar 0,171% , untuk peluang korban jiwa pada negara Singapura dengan berstatus dalam meninggal sebesar 0,079% dan peluang korban jiwa pada negara Singapura dengan berstatus dalam perawatan sebesar 0,750%

### 5.5.3 Membandingkan Nilai AIC

Kriteria kebaikan model dapat dilihat dari nilai AIC atau Akaike Information Criteria. Statistik AIC ini digunakan untuk mengukur seberapa dekat nilai harapan dan nilai observasi yang paling kecil. Untuk nilai AIC terkecil merupakan model terbaik.

Berikut adalah tabel perbandingan untuk model Loglinear dengan Logistik.

**Tabel 5.10.** Tabel perbandingan model Loglinear dengan Logistik

Model	AIC
Loglinear	1888,5
Logistik Multinomial	21145,36

Berdasarkan rumus estimasi prediksi frekuensi harapan masing-masing sel dari dua buah model dan berdasarkan nilai AIC untuk model loglinear dan logistik multinomial, didapatkan bahwa model loglinear memiliki nilai AIC terkecil dibandingkan dengan nilai logistik multinomial yaitu 1888,5. Maka disimpulkan bahwa model loglinear merupakan model terbaik.

## BAB VI

### PENUTUP

#### 6.1 Kesimpulan

Berdasarkan analisis yang telah peneliti lakukan maka didapatkan kesimpulan sebagai berikut:

1. Pada model Loglinear semua variabel dianggap sebagai variabel respon dan tidak ada perbedaan antara variabel respon dengan variabel penjelas dalam hal perlakuan, sedangkan pada model Logistik Multinomial variabel respon bergantung atau saling terikat dengan variabel penjelas.
2. Dengan menggunakan studi kasus yang sama yaitu menggunakan data untuk tabel kontigensi 2 arah. Didapatkan hasil bahwa model loglinear memberikan hasil dugaan dengan nilai yang relatif sama dengan model regresi logistik multinomial, yaitu pada nilai deviance atau Chi-Square pada kedua model didapatkan hasil yang sama sebesar 1650,52 sedangkan untuk nilai *Likelihood* ratio nya sebesar 1827,875. Walaupun model loglinear dengan model regresi logistik multinomial menunjukkan kesamaan, tetapi interpretasi terhadap model loglinear lebih kompleks karena melibatkan banyak variable, selain itu nilai AIC model loglinear lebih kecil dibandingkan model Logistik Multinomial

#### 6.2 Saran

Berdasarkan hasil penelitian yang dilakukan, saran yang dapat peneliti berikan sebagai berikut:

1. Hasil penelitian ini dapat menjadi salah satu acuan untuk melihat faktor apa saja yang saling berhubungan.
2. Taksiran parameter model dalam skripsi ini menggunakan metode maksimum *Likelihood*. Untuk lebih jelasnya bisa menggunakan metode penaksiran parameter lain salah satunya yaitu least squares.

3. Dikarenakan dalam penelitian ini sampel penelitian hanya menggunakan data kontigensi dua dimensi, maka peneliti menyarankan kepada peneliti selanjutnya untuk dapat meningkatkan seperti menggunakan data 3 dimensi atau 4 dimensi.

## DAFTAR PUSTAKA

- Susilo, Adityo dkk. 2020. *Coronavirus Disease 2019: Tinjauan Literatur Terkini*. Jurnal Penyakit Dalam Indonesia,07(01), :45-67.
- Alodokter. 2020. *Virus Corona*. <https://www.alodokter.com/virus-corona>. (24 Mei 2020)
- Agresti, Alan. 1984. *Analysis of Ordinal Categorical Data*. New York: John Jersey: John Wiley & Sons.
- Agresti, Alan. 1990. *Categorical Data Analysis*. USA: John Wiley and Sons.
- Hosmer, D.W dan Lemeshow, S. (1989). *Applied Logistic Regression*. New York: John Wiley dan Sons.
- Hosmer, D.W dan Lemeshow, S. (2000). *Applied Logistic Regression*. New York: John Wiley dan Sons.
- Dobson, A.J. 2001. *An Introduction to Generalized Linear Model*. Chapman Hall/CRC Texts in Statistical Science Series.
- Hapsari, G.S. 2011. *Model Loglinier untuk Tabel Kontigensi Tak Sempurna Berdimensi Tiga (Studi Kasus Jumlah Penduduk Kabupaten Sleman Tahun 2008 Menurut Umur, Pendidikan dan Jenis Kelamin)*. Fakultas MIPA UNY. Yogyakarta.
- Nugraha, J. (2017). *Pengantar Analisis Data Kategorik*. Yogyakarta: Deepublish.
- Nugraha, J. (2019). *Analisis Tabel Kontingensi menggunakan Model Loglinier*. Yogyakarta: Anggota IKAPI.
- Viestri, Dinar Ariana. 2015. *Pemodelan log linier dan regresi logistik biner bivariat pada medical check-up pegawai negeri sipil (PNS) Institut Teknologi Sepuluh Nopember (ITS)*. Undergraduate thesis, Institut Teknologi Sepuluh Nopember.
- Utomo, S. 2009. *Model Regresi Logistik untuk Menunjukkan Pengaruh Pendapatan Per Kapita, Tingkat Pendidikan, dan Status Pekerjaan terhadap Status Gizi Masyarakat Kota Surakarta*.



<https://www.scribd.com/doc/252113961/skripsi-model-regresi-logistik-pdf>.

(13 April 2020)

Maryana, 2013. *Model Log Linier yang Terbaik untuk Analisis Data Kualitatif pada Tabel Kontigensi Tiga Arah*.

<https://journal.unimal.ac.id/miej/article/view/77>. (13 April 2020)

Rosalia,dkk. 2011. *Analisis Model Log Linier untuk Mengetahui Kecenderungan Perilaku Anak Jalanan Binaan di Surabaya( Kasus Khusus Yayasan Arek Lintang-Alit*. Jurnal. Institut Teknologi Sepuluh Nopember.

McCullagh & Nelder, J. (1983). *Generalized Linear Models. 2nd Ed*. Chapman and Hall

## LAMPIRAN

**Lampiran 1** Data Jumlah Kasus Covid-19 Tahun 2020 Menurut Negara Indonesia Dan Singapura Berdasarkan Jenis Status Sembuh, Meninggal Dan Dalam Perawatan

Negara	Status			Total
	Sembuh	Meninggal	Perawatan	
Indonesia	1608	743	7032	9383
Singapura	1060	12	12552	13624
Total	2668	755	19584	23007

## Lampiran 2 Program Loglinear

```
covid = rbind(c(1608,743,7032), c(1060,12,12552))
test1 = chisq.test(covid,correct=F)
test1
try1 = loglin(covid,margin=list(1,2))
try1
try2 = loglin(covid,margin=list(1,2),fit=T,param=T)
try2
try3 = loglin(r covid,margin=list(c(1,2)),fit=T,param=T)
try3
log(sum(try2$fit[1,])/sum(try2$fit[2,]))/2
sum(covid); sum(try2$fit) # Both = n
G2 = 2 * sum(covid * log(covid /try2$fit))
G2
```

### Lampiran 3 Program Logistik Multinomial

```
datacovid<-read.csv(file="D:\\BISMILLAH          TA\\data          covid
20.csv",header=TRUE,sep=";")
datacovid
str(datacovid)
datacovid$status<-factor(datacovid$status)
datacovid$negara<-factor(datacovid$ntatus)
str(datacovid)
Model1 <- multinom(status ~ negara, data=covid19, trace=FALSE)
summary(Model1, cor=FALSE, Wald=TRUE)
```

#### Lampiran 4 Output Loglinear

```
> covid = rbind(c(1608,743,7032), c(1060,12,12552))
> test1 = chisq.test(covid,correct=F)
> test1

      Pearson's Chi-squared test

data: covid
X-squared = 1650.5, df = 2, p-value < 2.2e-16

> try1 = loglin(covid,margin=list(1,2))
2 iterations: deviation 0
> try1
$`lrt`
[1] 1827.717

$spearson
[1] 1650.521

$df
[1] 2

$margin
$margin[[1]]
[1] 1

$margin[[2]]
[1] 2

> try2 = loglin(covid,margin=list(1,2),fit=T,param=T)
2 iterations: deviation 0
> try2
$`lrt`
[1] 1827.717

$spearson
[1] 1650.521

$df
[1] 2

$margin
$margin[[1]]
[1] 1

$margin[[2]]
[1] 2

$fit
      [,1]      [,2]      [,3]
[1,] 1088.097 307.9135 7986.99
[2,] 1579.903 447.0865 11597.01
```

```

$param
$param$`(Intercept)`
[1] 7.422325

$param$`1`
[1] -0.1864667  0.1864667

$param$`2`
[1] -0.2436724 -1.5060390  1.7497114
> try3 = loglin(covid,margin=list(c(1,2)),fit=T,param=T)
2 iterations: deviation 0
> try3
$`lrt`
[1] 0

$spearson
[1] 0

$df
[1] 0

$margin
$margin[[1]]
[1] 1 2

$fit
      [,1] [,2] [,3]
[1,] 1608  743 7032
[2,] 1060  12 12552

$param
$param$`(Intercept)`
[1] 6.956706

$param$`1`
[1]  0.6605171 -0.6605171

$param$`2`
[1]  0.2176795 -2.4089045  2.1912250

$param$`1.2`
      [,1]      [,2]      [,3]
[1,] -0.452156  1.402378 -0.9502216
[2,]  0.452156 -1.402378  0.9502216
> log(sum(try2$fit[1,])/sum(try2$fit[2,]))/2
[1] -0.1864667
> sum(covid); sum(try2$fit) # Both = n
[1] 23007
[1] 23007
> G2 = 2 * sum(covid * log(covid/try2$fit))
> G2
[1] 1827.717

```

## Lampiran 5 Output Logistik Multinomial

```
> datacovid$status<-factor(datacovid$status)
> datacovid$negara<-factor(datacovid$status)
> str(datacovid)
'data.frame':  23007 obs. of  2 variables:
 $ negara: Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
 $ status: Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
> Model1 <- multinom(Status ~ Negara, data=covid19, trace=FALSE)

> summary(Model1, cor=FALSE, Wald=TRUE)
Call:
multinom(formula = Status ~ Negara, data = covid19, trace = FALSE)

Coefficients:
              (Intercept) Negara[T.Singapura]
Perawatan    2.2475394             4.704921
Sembuh       0.7720426             3.708826

Std. Errors:
              (Intercept) Negara[T.Singapura]
Perawatan    0.03857601             0.2913398
Sembuh       0.04435990             0.2936363

Value/SE (Wald statistics):
              (Intercept) Negara[T.Singapura]
Perawatan    58.26261             16.14925
Sembuh       17.40406             12.63068

Residual Deviance: 21137.36
AIC: 21145.36
```

**Lampiran 6** Tabel *Chi-Square*

<b>df</b>	<b>0.1</b>	<b>0.05</b>	<b>0.025</b>	<b>0.01</b>	<b>0.005</b>	<b>0.001</b>
1	2.706	3.841	5.024	6.635	7.879	10.828
2	4.605	5.991	7.378	9.21	10.597	13.816
3	6.251	7.815	9.348	11.345	12.838	16.266
4	7.779	9.488	11.143	13.277	14.86	18.467
5	9.236	11.07	12.833	15.086	16.75	20.515
6	10.645	12.592	14.449	16.812	18.548	22.458
7	12.017	14.067	16.013	18.475	20.278	24.322
8	13.362	15.507	17.535	20.09	21.955	26.124
9	14.684	16.919	19.023	21.666	23.589	27.877
10	15.987	18.307	20.483	23.209	25.188	29.588
11	17.275	19.675	21.92	24.725	26.757	31.264
12	18.549	21.026	23.337	26.217	28.3	32.909
13	19.812	22.362	24.736	27.688	29.819	34.528
14	21.064	23.685	26.119	29.141	31.319	36.123
15	22.307	24.996	27.488	30.578	32.801	37.697
16	23.542	26.296	28.845	32	34.267	39.252
17	24.769	27.587	30.191	33.409	35.718	40.79
18	25.989	28.869	31.526	34.805	37.156	42.312
19	27.204	30.144	32.852	36.191	38.582	43.82
20	28.412	31.41	34.17	37.566	39.997	45.315
21	29.615	32.671	35.479	38.932	41.401	46.797
22	30.813	33.924	36.781	40.289	42.796	48.268
23	32.007	35.172	38.076	41.638	44.181	49.728
24	33.196	36.415	39.364	42.98	45.559	51.179
25	34.382	37.652	40.646	44.314	46.928	52.62
26	35.563	38.885	41.923	45.642	48.29	54.052
27	36.741	40.113	43.195	46.963	49.645	55.476
28	37.916	41.337	44.461	48.278	50.993	56.892
29	39.087	42.557	45.722	49.588	52.336	58.301
30	40.256	43.773	46.979	50.892	53.672	59.703
31	41.422	44.985	48.232	52.191	55.003	61.098
32	42.585	46.194	49.48	53.486	56.328	62.487
33	43.745	47.4	50.725	54.776	57.648	63.87
34	44.903	48.602	51.966	56.061	58.964	65.247
35	46.059	49.802	53.203	57.342	60.275	66.619
36	47.212	50.998	54.437	58.619	61.581	67.985
37	48.363	52.192	55.668	59.893	62.883	69.346
38	49.513	53.384	56.896	61.162	64.181	70.703
39	50.66	54.572	58.12	62.428	65.476	72.055



---

<b>40</b>	51.805	55.758	59.342	63.691	66.766	73.402
<b>41</b>	52.949	56.942	60.561	64.95	68.053	74.745
<b>42</b>	54.09	58.124	61.777	66.206	69.336	76.084
<b>43</b>	55.23	59.304	62.99	67.459	70.616	77.419
<b>44</b>	56.369	60.481	64.201	68.71	71.893	78.75
<b>45</b>	57.505	61.656	65.41	69.957	73.166	80.077
<b>46</b>	58.641	62.83	66.617	71.201	74.437	81.4
<b>47</b>	59.774	64.001	67.821	72.443	75.704	82.72
<b>48</b>	60.907	65.171	69.023	73.683	76.969	84.037
<b>49</b>	62.038	66.339	70.222	74.919	78.231	85.351
<b>50</b>	63.167	67.505	71.42	76.154	79.49	86.661
<b>60</b>	74.397	79.082	83.298	88.379	91.952	99.607
<b>70</b>	85.527	90.531	95.023	100.425	104.215	112.317
<b>80</b>	96.578	101.879	106.629	112.329	116.321	124.839
<b>90</b>	107.565	113.145	118.136	124.116	128.299	137.208
<b>100</b>	118.498	124.342	129.561	135.807	140.169	149.449

---