PEMODELAN *BIVARIATE POISSON LOG-NORMAL REGRESSION* PADA JUMLAH KASUS KUSTA DI INDONESIA

SALSABILA RAHMADHANI S.
H051201068

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

PEMODELAN *BIVARIATE POISSON LOG-NORMAL REGRESSION* PADA JUMLAH KASUS KUSTA DI INDONESIA

SALSABILA RAHMADHANI S.
H051201068

PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

PEMODELAN *BIVARIATE POISSON LOG-NORMAL REGRESSION* PADA JUMLAH KASUS KUSTA DI INDONESIA

SALSABILA RAHMADHANI S.
H051201068

Skripsi

sebagai salah satu syarat untuk mencapai gelar sarjana

Program Studi Statistika

pada

DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

SKRIPSI

PEMODELAN *BIVARIATE POISSON LOG-NORMAL REGRESSION* PADA JUMLAH KASUS KUSTA DI INDONESIA

SALSABILA RAHMADHANI S.
H051201068

Skripsi,

telah dipertahankan di depan Panitia Ujian Sarjana pada tanggal 16 Agustus 2024 dan dinyatakan telah memenuhi syarat kelulusan pada

Program Studi Statistika
Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan:	Mengetahui:
Pembimbing tugas akhir,	Ketua Program Studi,
Siswanto, S.Si., M.Si.	Dr. Anna Islamiyati, S.Si., M.Si.
NIP. 19920107 201903 1 012	NIP. 19770808 200501 2 002

KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS HASANUDDIN
FAKULTAS MIPA

PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Pemodelan *Bivariate Poisson Log-Normal Regression* pada Jumlah Kasus Kusta di Indonesia" adalah benar karya saya dengan arahan dari pembimbing Siswanto, S.Si., M.Si. sebagai Pembimbing Utama. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas pembuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 16 Agustus 2024

Salsabila Rahmadhani S.
NIM. H051201068

UCAPAN TERIMA KASIH

Dengan menyebut nama Allah yang Maha Pengasih, Maha Penyayang. Puji syukur penulis panjatkan kepada Allah *Subhanahu Wa Ta'ala* atas segala limpahan cinta-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "**Pemodelan *Bivariate Poisson Log-Normal Regression* pada Kasus Kusta di Indonesia**". Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa Sallam* beserta keluarga dan para sahabatnya.

Penulis menyampaikan ucapan terima kasih kepada **Bapak Siswanto, S.Si., M.Si.** selaku Pembimbing Utama yang dengan penuh kesabaran telah meluangkan waktu, pemikiran, semangat, dan bimbingan kepada penulis selama proses penulisan tugas akhir ini. Terima kasih kepada **Ibu Prof. Dr. Dr. Georgina Tinungki, S.Si., M.Si.** dan Ibu **Dr. Erna Tri Herdiani, S.Si., M.Si.** selaku Tim Penguji yang senantiasa memberikan saran dan kritikan yang membangun dalam penulisan tugas akhir ini. Terima kasih kepada **Pimpinan Universitas Hasanuddin, Departemen Statistika, Jajaran Dosen,** dan **Staf Departemen Statistika** yang telah memfasilitasi, memberikan ilmu bermanfaat, dan membantu penulis selama menempuh studi.

Penghargaan dan ucapan terima kasih yang tulus juga penulis ucapkan kepada kedua orang tua penulis yang tercinta, Ayahanda **Setiawan Azis, S.STP, MSP** dan Ibunda **Titin Florentina, M.Psi., Psikolog** kepada nenek penulis **Eyang Tatik Oding** beserta **keluarga besar** penulis yang telah memberikan dukungan tanpa syarat, semangat, dan doa yang senantiasa mengiringi langkah penulis. Terima kasih kepada saudara penulis **Alm. Muhammad Thufeil** yang senantiasa membawa kebahagiaan di dalam mimpi penulis. Terima kasih kepada **Ridwan, Ayuni, Isra, Divia, Isti, Fadlan, Mukhlis, Hakam, Ryan, Fahmi, Reza, Theo, dan Faiqha** yang sudah menjadi garda terdepan dalam proses pembelajaran dan penyelesaian tugas akhir ini. Terima kasih kepada **Faldi, Nahdi, Mary, Izzul, Edward, Heri, Taufiq, Razy, Febi, Najlah, Ayu Afri, Aliyah, Ngkal dan** teman-teman **POIS20N** yang namanya tidak bisa disebut satu persatu yang telah menemani penulis menghadapi suka duka selama perkuliahan. Terima kasih kepada teman-teman **Statistika'20** yang semangat, kecerdasan dan ambisnya dalam belajar membuat penulis tertekan dan cepat dalam menyelesaikan skripsi ini. Terima kasih kepada sahabat **Markaz Shohibul Qur'an** yang cintanya mengangkat inspirasi penulis. Terima kasih kepada keluarga besar **Himastat FMIPA Unhas** atas kebersamaannya selama ini. Terima kasih kepada **Meganthropus Family** yang selalu memberikan doa dan dukungan dimanapun penulis berada.

Penulis telah berusaha semaksimal mungkin, namun menyadari bahwa masih banyak kekurangan dalam penyusunan tugas akhir ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini dapat memberikan manfaat untuk berbagai pihak.

Makassar, 16 Agustus 2024

Penulis,

Salsabila Rahmadhani S.

ABSTRAK

Salsabila Rahmadhani S. **Pemodelan *Bivariate Poisson Log-Normal Regression* Pada Jumlah kasus Kusta di Indonesia** (dibimbing oleh Siswanto).

Latar belakang. Dua variabel respon berupa data diskrit yang saling berkorelasi dapat dimodelkan menggunakan regresi *Bivariate Poisson*. Dalam regresi *Poisson* sering terjadi pelanggaran asumsi yaitu overdispersi yang dapat diatasi dengan model *mixed Poisson* seperti *Poisson Log-Normal*. Salah satu data diskrit yang diasumsikan berdistribusi *Poisson* adalah jumlah kasus kusta yang diklasifikasi/kan oleh *World Health Organization* (WHO) menjadi dua tipe yaitu *Pauci Bacillary* (PB) dan *Multi Bacillary* (MB). Kedua variabel tersebut saling berkorelasi dan terjadi kasus overdispersi sehingga dimodelkan dengan *Bivariate Poisson Lognormal Regression* (BPLNR). **Tujuan.** Penelitian ini bertujuan untuk mendapatkan estimasi parameter dan mengetahui faktor faktor yang diduga memengaruhi jumlah kasus kusta PB dan MB di Indonesia tahun 2021 dengan model BPLNR. **Metode.** Estimasi parameter BPLNR menggunakan *Maximum Likelihood Estimation* (MLE) dengan algoritma *Newton Raphson* dan pengujian hipotesis menggunakan metode *Maximum Likelihood Ratio Test* (MLRT). **Hasil.** BPLNR dapat memodelkan data jumlah kasus kusta PB dan MB di Indonesia tahun 2021 serta mengatasi overdispersi dengan parameter dispersi sebesar $1{,}001648$. Faktor-faktor yang memengaruhi jumlah kasus kusta PB dan MB di Indonesia tahun 2021 yaitu, kepadatan penduduk, persentase penduduk miskin, persentase rumah tangga dengan akses sanitasi layak, jumlah fasilitas kesehatan, jumlah tenaga medis, dan persentase rumah tangga dengan akses air minum layak. **Kesimpulan.** Dengan demikian, dapat disimpulkan bahwa metode BPLNR dapat memodelkan data jumlah kasus kusta PB dan MB di Indonesia tahun 2021, dan mendapatkan faktor-faktor yang berpengaruh secara signifikan terhadap jumlah kasus kusta serta parameter dispersi yang dapat mengakomodasi overdispersi.

Kata kunci: *Bivariate Poisson Log-Normal Regression*, Kusta PB, Kusta MB, *Maximum Likelihood Estimation*, Overdispersi.

ABSTRACT

Salsabila Rahmadhani S. **Bivariate Poisson Log-Normal Regression Modeling on the Number of Leprosy Cases in Indonesia** (supervised by Siswanto).

Background. Two response variables in the form of correlated discrete data can be modeled using Bivariate Poisson regression. In Poisson regression, there is often a violation of assumptions, namely overdispersion, which can be overcome by mixed Poisson models such as Log-Normal Poisson. One of the discrete data assumed to be Poisson distributed is the number of leprosy cases classified by the World Health Organization (WHO) into two types, namely Pauci Bacillary (PB) and Multi Bacillary (MB). The two variables are correlated and there is a case of overdispersion so that it is modeled with Bivariate Poisson Lognormal Regression (BPLNR). **Aim**. This study aims to obtain parameter estimates and determine the factors that are thought to affect the number of PB and MB leprosy cases in Indonesia in 2021 with the BPLNR model. **Methods.** BPLNR parameter estimation uses Maximum Likelihood Estimation (MLE) with the Newton Raphson algorithm and hypothesis testing using the Maximum Likelihood Ratio Test (MLRT) method. **Results**. BPLNR can model data on the number of PB and MB leprosy cases in Indonesia in 2021 and overcome overdispersion with a dispersion parameter of $1{,}001648$. Factors that influence the number of PB and MB leprosy cases in Indonesia in 2021 are population density, percentage of poor people, percentage of households with access to proper sanitation, number of health facilities, number of medical personnel, and percentage of households with access to proper drinking water. **Conclusion**. Thus, it can be concluded that the BPLNR method can model data on the number of PB and MB leprosy cases in Indonesia in 2021, and obtain factors that significantly affect the number of leprosy cases and dispersion parameters that can accommodate overdispersion.

Keywords: Bivariate Poisson Log-Normal Regression, Leprosy PB, Leprosy MB, Maximum Likelihood Estimation, Overdispersion.

DAFTAR ISTILAH

Istilah	Arti dan Penjelasan
Analitik	Metode penyelesaian model matematika dengan rumus-rumus aljabar yang sudah baku (lazim).
Demografi	Ilmu mengenai susunan, jumlah, dan perkembangan penduduk
Diskrit	Bilangan bulat, tidak bisa dipecah menjadi angka pecahan atau desimal
Distribusi	Penyebaran sesuatu secara merata
Eksplisit	Dinyatakan dengan jelas dan tegas, tanpa ada keraguan atau implikasi yang tersembunyi.
Estimator	Statistik sampel yang digunakan untuk menduga suatu parameter
Fungsi Massa Peluang	Fungsi yang menggambarkan probabilitas bahwa variabel acak diskrit akan mengambil nilai tertentu.
Fungsi Kepadatan Peluang	Fungsi yang menggambarkan probabilitas relatif dari variabel acak kontinu, dikenal sebagai fungsi kepadatan probabilitas (PDF).
Gradien	Vektor yang berisi turunan pertama dari fungsi, menunjukkan arah laju perubahan tercepat.
Hipotesis	Pernyataan atau dugaan sementara yang bisa diuji melalui eksperimen atau observasi untuk menentukan validitasnya.
Homogenitas	Kondisi dimana tidak terdapat perbedaan variansi
Implisit	Sesuatu yang tersirat, atau tidak dinyatakan secara langsung.
Iterasi	Proses pengulangan langkah-langkah dalam suatu algoritma atau prosedur sampai kondisi tertentu tercapai.
Koefisien	Terjadi atau dilakukan pada saat yang sama.
Kontinu	Mengacu pada data atau variabel yang dapat mengambil nilai apa pun dalam suatu rentang tertentu, tidak terbatas pada nilai diskrit.
Konvergen	Kondisi ketika suatu urutan atau seri mendekati nilai tertentu seiring berjalannya waktu atau iterasi.
Korelasi	Ukuran statistik yang menunjukkan sejauh mana dua variabel berhubungan atau bergerak bersama.
Lesi Kulit	Kerusakan atau abnormalitas pada kulit yang bisa berupa bintik, luka, atau pertumbuhan abnormal.
Likelihood	Fungsi yang mengukur seberapa baik model statistik menjelaskan data yang diamati, sering digunakan dalam estimasi parameter.

Istilah	Arti dan Penjelasan
Logaritma Natural	Logaritma berbasis e (sekitar 2.718), sering dilambangkan sebagai ln(x).
Matriks	Susunan bilangan dalam bentuk baris dan kolom yang digunakan dalam berbagai operasi matematika.
Overestimate	Menaksir atau memperkirakan sesuatu lebih tinggi dari nilai sebenarnya.
Parameter	Nilai yang memperkirakan nilai populasi melalui suatu fungsi atau persamaan model
Parsial	Bagian dari keseluruhan, sering digunakan dalam konteks turunan parsial dalam kalkulus, yang mengukur perubahan fungsi terhadap satu variabel saat variabel lain tetap.
Pemodelan	Proses membuat representasi matematika dari sistem atau proses untuk analisis dan prediksi.
Prediktor	Variabel yang digunakan untuk memprediksi atau menjelaskan variabel respon; juga dikenal sebagai variabel independen.
Probabilitas	Sinonim dari peluang, ukuran kemungkinan terjadinya sesuatu peristiwa, biasanya dinyatakan dalam nilai 0 dan 1.
Respon	Variabel yang diukur atau diamati untuk melihat pengaruh dari variabel prediktor dalam sebuah studi atau eksperimen; juga dikenal sebagai variabel dependen.
Simultan	Terjadi atau dilakukan pada saat yang sama.
Skewness	Ukuran ketidaksimetrisan distribusi probabilitas suatu variabel acak; positif jika distribusi condong ke kanan, negatif jika condong ke kiri.
Variabel Acak	Variabel yang hasilnya merupakan nilai acak dan ditentukan oleh probabilitas.
Variansi	Ukuran penyebaran atau dispersi dari sekumpulan data; rata-rata kuadrat deviasi setiap nilai dari rata-rata.
Vektor	Objek matematika yang memiliki besaran dan arah, sering digunakan dalam aljabar linier dan statistik.

DAFTAR LAMBANG DAN SINGKATAN

Lambang/singkatan	Arti dan Penjelasan
α	“alpha” taraf signifikansi
β	“beta” parameter regresi
θ	“theta” parameter
τ	“tau” parameter dispersi
μ	“mu” parameter distribusi *Poisson*
λ, τ^2	“lambda” parameter distribusi *Log-Normal*
π	“pi” bilangan konstanta
∂	*“delta” turunan*
ω,Ω	“Omega” himpunan parameter
x	Variabel prediktor
Z	Variabel acak berdistribusi *Poisson*
$v.V$	Variabel acak berdistribusi *Log-Normal*
y, Y	Variabel respon
H_0	Hipotesis nol
H_1	Hipotesis alternatif
j	Jumlah variabel respon
k	Jumlah variabel prediktor
i	Jumlah pengamatan
n	Jumlah pengamatan
T	*Transpose*
r	Bilangan yang membuat nilai non-negatif
L	Fungsi *Likelihood*
l	Fungsi ln *likelihood*
s	Iterasi pada algoritma *Newton Raphson*
t	Statistik uji t hitung
D	*Deviance*
H	*Hessian*
g	Gradien
χ^2	*Chi-Square*
$Cov(.)$	kovariansi
$\exp(.)$	Eksponen
$E(.)$	Ekspektasi
$Var(.)$	Variansi
$\ln(\cdot)$	Logaritma Natural
$f(y)$	Fungsi distribusi
$\min(.)$	Minimum
$F(x)$	Fungsi distribusi kumulatif
$F_0(x)$	Fungsi probabilitas
$g(v)$	Fungsi distribusi *Log-Normal*
r_{y_1,y_2}	Korelasi variabel respon

Lambang/singkatan	Arti dan Penjelasan
!	Faktorial
$p - value$	Nilai probabilitas
BP	*Bivariate Poisson*
BPLNR	*Bivariate Poisson Log-Normal*
LN	*Log-Normal*
MB	*Multi Bacillary*
MLE	*Maximum Likelihood Estimation*
MLRT	*Maximum Likelihood Ratio Test*
PB	*Pauci Bacillary*
VIF	*Variance Inflation Factor*
WHO	*World Health Organization*
WER	*Weekly Epidemiological Record*

DAFTAR ISI

Halaman

DAFTAR TABEL

Nomor Urut **Halaman**

DAFTAR GAMBAR

Nomor Urut **Halaman**

DAFTAR LAMPIRAN

Nomor Urut **Halaman**

BAB I
PENDAHULUAN

1.1. Latar Belakang

Regresi *Poisson* adalah sebuah metode yang bertujuan untuk memodelkan variabel respon berupa data diskrit dan berdistribusi *Poisson*. Data yang berdistribusi Poisson merupakan data yang memiliki probabilitas kecil dan bergantung pada interval waktu atau daerah tertentu, dengan pengamatan berupa variabel diskrit dan antar variabel saling bebas (Arkandi & Winahju, 2016). Variabel respon diskrit berdistribusi *Poisson* dapat dimodelkan dengan pendekatan regresi *Poisson*, sedangkan dua data diskrit yang berdistribusi *Poisson* dan saling berkorelasi dapat dimodelkan dengan regresi *Bivariate Poisson*. Penelitian terkait model regresi *Bivariate Poisson* dilakukan oleh Kaombe (2024) pada data pendidikan dalam satu tahun dan jumlah anak yang dilahirkan. Selain itu, Prahutama dkk. (2021) juga menggunakan model regresi *Bivariate Poisson* pada data kematian ibu dan bayi di Jawa Tengah. Penelitian tersebut memodelkan data diskrit, namun mengabaikan asumsi dalam regresi *Poisson,* yaitu equidispersi.

Equidispersi adalah kondisi data yang memiliki nilai *mean* dan variansi yang sama. Namun dalam banyak kasus, asumsi equidispersi tidak terpenuhi. Beberapa kasus yang sering ditemui memiliki kondisi data yang nilai variansi lebih besar daripada nilai *mean*-nya, atau disebut sebagai overdispersi (McCullagh & Nelder, 1989). Masalah overdispersi dapat menyebabkan *underestimate* pada standar *error* estimasi parameter yang mengakibatkan kesalahan dalam pengambilan kesimpulan (Hilbe, 2014). Salah satu metode yang dapat digunakan untuk mengatasi masalah overdispersi adalah dengan menggunakan distribusi *mixed Poisson,* yaitu campuran distribusi *Poisson* dengan suatu distribusi kontinu. Salah satu distribusi *mixed Poisson* adalah distribusi *Poisson Log-Normal* yang merupakan campuran antara distribusi *Poisson* dan *Log-Normal*. Distribusi *Log-Normal* merupakan distribusi yang jika di transformasikan dengan logaritma natural akan menjadi distribusi normal (Walpole dkk., 2016). Novkaniza dkk. (2023) telah melakukan penelitian menggunakan model *Poisson Log-Normal* pada data tingkat buta huruf di Provinsi Kepulauan Riau. Penelitian tersebut menghasilkan model yang dapat mengatasi masalah overdispersi. Shirazi & Lord (2019) membandingkan model *Poisson-gamma* dan *Poisson Log-Normal* menggunakan dua data set, yaitu data kecelakaan kendaraan tunggal yang terjadi di jalan raya dua jalur pedesaan Michigan pada tahun 2006 dan kecelakaan yang terjadi antara 2012 dan 2014 di arteri empat jalur perkotaan Texas. Hasil Penelitian menunjukkan bahwa untuk kedua data set tersebut, model *Poisson Log-Normal* lebih baik untuk memodelkan data tersebut.

Salah satu data yang diasumsikan berdistribusi *Poisson* adalah data jumlah kasus kusta yang merupakan penyakit kronis akibat infeksi bakteri *Mycobacterium leprae*. Terdapat dua jenis penyakit kusta, yaitu *Pauci Bacillary* (PB) yang ditandai dengan adanya lima atau lebih sedikit lesi kulit tanpa bakteri yang terlihat pada pemeriksaan mikroskopis dan *Multi Bacillary* (MB) yang ditandai dengan lebih dari lima lesi kulit dan kehadiran bakteri yang dapat terlihat pada pemeriksaan mikroskopis (Kementerian Kesehatan RI, 2021). Berdasarkan laporan dari *World*

Health Organization (WHO) dalam *Weekly Epidemiological Record* (WER) tahun 2021, Indonesia berada di posisi ketiga sebagai negara dengan jumlah kasus kusta terbanyak di dunia, setelah India dan Brazil. Pada tahun 2021, Indonesia melaporkan 10.876 kasus baru kusta. Penyakit kusta dipengaruhi oleh beberapa aspek sosial ekonomi, aspek lingkungan, aspek demografi, dan aspek perilaku (Jariwala dkk., 2013). Selain itu, aspek fasilitas kesehatan juga memengaruhi terjadinya kusta (Kerr-Pontes dkk., 2006). Pada aspek sosial ekonomi meliputi penduduk miskin dan rumah tangga dengan alas lantai tanah. Aspek lingkungan meliputi rumah sehat, dan jamban sehat. Aspek demografi meliputi kepadatan penduduk. Aspek perilaku meliputi rumah tangga ber-PHBS, dan sarana air bersih. Sedangkan aspek fasilitas kesehatan meliputi banyak *public health*, dan jumlah tenaga medis.

Data jumlah kasus kusta di Indonesia tahun 2021 merupakan data diskrit yang memiliki probabilitas kejadian yang kecil, sehingga data berdistribusi *Poisson*. Data kusta PB dan MB memiliki variabel dependen yang saling berkorelasi, sehingga model *Bivariate Poisson* dapat diterapkan. Selain itu, data kusta tahun 2021 mengalami overdispersi yang dapat menghasilkan kesimpulan yang tidak valid. Untuk mengatasi masalah ini, digunakan metode *Poisson Log-Normal* dengan memperkenalkan distribusi *Log-Normal* pada parameter model. Dengan menggabungkan pendekatan ini, dikembangkan metode *Bivariate Poisson Log-Normal Regression*. Metode penaksiran parameter model *Bivariate Poisson Log-Normal Regression* (BPLNR) yang digunakan yaitu metode *Maximum Likelihood Estimastion* (MLE) dengan memaksimumkan fungsi *likelihood* dari distribusi data. Namun beberapa fungsi *likelihood* tidak dapat diselesaikan secara analitik menggunakan metode MLE karena diperoleh bentuk implisit, sehingga harus diselesaikan secara iteratif. Salah satu metode iterasi yang dapat digunakan adalah *Newton Raphson*. Dengan latar belakang tersebut, penelitian ini bertujuan untuk mengaplikasikan metode BPLNR dalam analisis data kusta PB dan MB di Indonesia tahun 2021.

1.2. Batasan Masalah

Batasan masalah yang dibahas pada penelitian ini adalah:

1. Data yang digunakan adalah data jumlah kusta tahun 2021 dengan ruang lingkup penelitian adalah semua Provinsi di Indonesia.
2. Metode estimasi parameter model *Bivariate Poisson Log-Normal Regression* menggunakan *Maximum Likelihood Estimation* (MLE) dengan algoritma *Newton Raphson*

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, tujuan dari penelitian ini adalah:

1. Memperoleh estimasi parameter model *Bivariate Poisson Log-Normal Regression.*
2. Memperoleh faktor-faktor yang memengaruhi jumlah kasus kusta di Indonesia menggunakan model *Bivariate Poisson Log-Normal Regression.*

1.4. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

1. Dapat memperluas pemahaman teoritis dan praktis terkait model *Bivariate Poisson Log-Normal Regression* terutama dalam estimasi parameter.
2. Penerapan model *Bivariate Poisson Log-Normal Regression* dalam bidang kesehatan sebagai metode akternatif dalam mengetahui faktor-faktor apa saja yang memengaruhi jumlah kasus kusta PB dan MB di Indonesia pada tahun 2021.
3. Dapat menjadi bahan informasi bagi masyarakat maupun instansi pemerintah dalam mengevaluasi upaya penurunan jumlah kasus kusta.

1.5. Teori

1.5.1. Regresi *Poisson*

Regresi *Poisson* adalah model regresi non-linier yang bertujuan untuk memodelkan variabel respon berupa data diskrit dan berdistribusi *Poisson*. Distribusi *Poisson* adalah distribusi probabilitas yang menggambarkan kejadian-kejadian yang memiliki probabilitas kejadian yang kecil, dimana kejadian tersebut bergantung pada interval waktu tertentu atau terjadi di suatu daerah tertentu. Jika variabel acak diskrit (Y) berdistribusi *Poisson* dengan parameter $\mu > 0$, (Arkandi & Winahju, 2016). Maka fungsi massa peluang dinyatakan pada Persamaan (1).

$$f(y;\mu) = \frac{\exp(-\mu)\,\mu^y}{y!}, y = 0, 1, 2 \ldots \tag{1}$$

Distribusi *Poisson* memiliki nilai *mean* dan variansi yang sama, yaitu $E(Y) = Var(Y) = \mu$ (Lukman dkk., 2023). Persamaan *mean* dan variansi ini disebut equidispersi yang menjadi asumsi dalam pemodelan regresi *Poisson*. Model regresi *Poisson* dengan parameter μ yang dihubungkan dengan variabel prediktor menggunakan fungsi penghubung $\ln(\cdot)$ adalah pada Persamaan (2) (Cameron & Trivedi, 2013).

$$\begin{aligned} g(\mu_i) &= \ln(\mu_i) \\ \ln(\mu_i) &= (\boldsymbol{x}_i^T \boldsymbol{\beta}) \\ \exp(\ln(\mu_i)) &= \exp(\boldsymbol{x}_i^T \boldsymbol{\beta}) \\ \mu_i &= \exp(\boldsymbol{x}_i^T \boldsymbol{\beta}) \end{aligned} \tag{2}$$

dengan

μ_i : rata-rata variabel respon yang berdistribusi *Poisson*.

$\boldsymbol{x}_i^T$: $[1 \quad x_{i1} \quad x_{i2} \quad \ldots \quad x_{ik}]$ menunjukkan vektor yang berukuran $1 \times (k+1)$ dari variabel prediktor.

$\boldsymbol{\beta}$: $[\beta_0 \quad \beta_1 \quad \beta_2 \quad \ldots \quad \beta_k]^T$ menunjukkan vektor yang berukuran $(k+1) \times 1$ dari parameter regresi.

1.5.2. Regresi *Bivariate Poisson*

Regresi *Bivariate Poisson* adalah metode yang digunakan untuk memodelkan dua variabel respon diskrit yang berdistribusi *Poisson* dan saling berkorelasi (Karlis & Ntzoufras, 2005). Jika variabel acak diskrit $Y_1 \,\mathrm{dan}\, Y_2$ saling berkorelasi dan berdistribusi Poisson dengan parameter $\mu_j > 0$ maka fungsi massa peluang dinyatakan pada Persamaan (3) (Tzougas & di Cerchiara, 2023).

$$f\left(y_j; \mu_j; j = 1,2\right) = \frac{\exp(-\mu_j)\, \mu_j^{y_j}}{y_j!}, y_j = 0, 1, 2 \ldots \tag{3}$$

Distribusi *Bivariate Poisson* terjadi saat variabel acak $Z_0, Z_1, \mathrm{dan}\, Z_2$ masing-masing berdistribusi *Poisson* dengan parameter $\mu_0, \mu_1,$ dan μ_2. Terdapat variabel acak $Y_1 \,\mathrm{dan}\, Y_2$ yang terbentuk dari variabel $Z_0, Z_1, \mathrm{dan}\, Z_2$ yang saling bebas seperti pada Persamaan (4).

$$\begin{aligned} Y_1 &= Z_1 + Z_0 \\ Y_2 &= Z_2 + Z_0 \end{aligned} \tag{4}$$

Oleh karena itu, variabel acak Y_1 dan Y_2 secara bersama-sama berdistribusi *Bivariate Poisson* dengan fungsi massa peluang seperti pada Persamaan (5).

$$f_{BP}(y_1, y_2) = \exp(-\mu) \sum_{r=0}^{\min\{y_1, y_2\}} \frac{\mu_1^{y_1 - r}\, \mu_2^{y_2 - r}\, \mu_0^r}{(y_1 - r)!\,(y_2 - r)!\, r!}; y_1, y_2 \geq 0 \tag{5}$$

dimana $\mu = \mu_1 + \mu_2 + \mu_0$ dan r merupakan bilangan yang membuat nilai non-negatif bagi $y_1 - r; y_2 - r; r$ (Wang, 2017). Nilai *mean* dan variansi dari variabel respon $Y_1 \,\mathrm{dan}\, Y_2$ bernilai sama yaitu $E(Y_1) = Var(Y_1) = \mu_1 + \mu_0$ dan $E(Y_2) = Var(Y_2) = \mu_2 + \mu_0$.

Parameter μ_j dihubungkan dengan variabel prediktor menggunakan fungsi penghubung $\ln(\cdot)$. Sehingga model regresi *Bivariate Poisson* seperti pada Persamaan (6) (Kocherlakota & Kocherlakota, 2007).

$$\begin{aligned} \ln(\mu_{ij} + \mu_0) &= (\boldsymbol{x}_i^T \boldsymbol{\beta}_j) \\ \mu_{ij} + \mu_0 &= \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_j\right); \; j = 1,2 \end{aligned} \tag{6}$$

dengan:

μ_{ij}	:	rata-rata variabel respon ke$-j$ yang berdistribusi *Poisson*.
$\boldsymbol{x}_i^T$	:	$[1 \quad x_{i1} \quad x_{i2} \quad \ldots \quad x_{ik}]$ menunjukkan vektor yang berukuran $1 \times (k+1)$ dari variabel prediktor.
$\boldsymbol{\beta}_j$	:	$[\beta_{0j} \quad \beta_{1j} \quad \beta_{2j} \quad \ldots \quad \beta_{kj}]^T$ menunjukkan vektor yang berukuran $(k+1) \times 1$ dari parameter regresi.

Pengujian distribusi *Bivariate Poisson* dilakukan dengan menggunakan pendekatan *index of dispersion test.* Pengujian tersebut dilakukan untuk mengidentifikasi variabel respon Y_1 dan Y_2 mengikuti distribusi *Bivariate Poisson* atau tidak. Pengujian distribusi *Bivariate Poisson* menggunakan pendekatan *index of* *dispersion test* dilakukan dengan tahapan sebagai berikut (Best, 1999).

a. Hipotesis

H_0 : $F(x) = F_0(x)$ untuk Y_1 dan Y_2 (variabel respon Y_1 dan Y_2 mengikuti distribusi *Bivariate Poisson*).

H_1 : $F(x) \neq F_0(x)$ untuk Y_1 dan Y_2 (variabel respon Y_1 dan Y_2 tidak mengikuti distribusi *Bivariate Poisson*).

b. Statistik Uji

Pengujian distribusi *Bivariate Poisson* menggunakan pendekatan *index of dispersion test* berdasarkan Persamaan (7).

$$I_B = \frac{n\left(\bar{Y}_2 S_{Y_1}^2 - 2m_{11}^2 + \bar{Y}_1 S_{Y_2}^2\right)}{(\bar{Y}_1\bar{Y}_2 - m_{11}^2)} \sim \chi^2(2n-3) \tag{7}$$

dengan $S_{Y_1}^2 = \frac{\sum_{i=1}^{n}(Y_{i1}-\bar{Y}_1)^2}{n}$, $S_{Y_2}^2 = \frac{\sum_{i=1}^{n}(Y_{i2}-\bar{Y}_2)^2}{n}$, $m_{11} = \frac{\sum_{i=1}^{n}(Y_{i1}-\bar{Y}_1)(Y_{i2}-\bar{Y}_2)}{n}$, Y adalah variabel respon dan n adalah jumlah sampel.

c. Kriteria Uji

Menolak H_0 jika $|I_B| > \chi^2_{(\alpha;2n-3)}$ pada taraf signifikansi α.

1.5.3. Equidispersi

Salah satu asumsi yang harus dipenuhi dalam regresi *Poisson* adalah asumsi equidispersi, yaitu kondisi di mana nilai *mean* dan variansi dari variabel respon adalah sama. Namun dalam banyak kasus, asumsi ini sering sekali tidak terpenuhi. Jika variansi dari data lebih besar dari nilai *mean*-nya pada model *Poisson* ($Var(Y) > E(Y)$), maka dapat dikatakan bahwa data tersebut mengalami overdispersi (McCullagh & Nelder, 1989). Beberapa penyebab overdispersi diantaranya yaitu, korelasi positif antar variabel respon, variasi antar variabel respon yang berlebihan, pelanggaran asumsi distribusi dari data, kejadian sebelumnya memengaruhi kejadian berikutnya, model menghilangkan variabel prediktor yang penting, terdapat *outlier* pada data, kesalahan fungsi penghubung, dan kesalahan pada asumsi hubungan linier antara variabel respon (Hilbe, 2014).

Metode untuk mendeteksi overdispersi dapat menggunakan uji *deviance*. Adapun tahapan uji overdispersi adalah sebagai berikut.

a. Hipotesis

H_0 : Tidak terdapat overdispersi pada model regresi *Poisson*.

H_1 : Terdapat overdispersi pada model regresi *Poisson*.

b. Statistik Uji

Pengujian overdispersi dilakukan dengan uji *Deviance* pada Persamaan (8).

$$D = 2\sum_{i=1}^{n}\left(\left(y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - \hat{\mu}_i)\right)\right) \tag{8}$$

dengan y adalah variabel respon, dan μ adalah rata-rata variabel respon.

c. Kriteria Uji

Kriteria penolakan H_0 adalah jika nilai $D/(n-k) > 1$ (Hilbe, 2014).

1.5.4. Regresi *Poisson Log-Normal*

Regresi *Poisson Log-Normal* adalah metode yang digunakan untuk memodelkan data diskrit yang mengalami overdispersi dari distribusi *Poisson Log-Normal*. Distribusi *Poisson Log-Normal* merupakan campuran distribusi *Poisson* dengan *Log-Normal* yang merupakan salah satu alternatif distribusi *mixed Poisson*. Distribusi *Log-Normal* pertama kali diperkenalkan oleh dua matematikawan Inggris, Franics Galton dan Donald McAlister pada tahun 1879. Distribusi *Log-Normal* merupakan distribusi yang jika di transformasikan dengan logaritma natural akan menjadi distribusi normal (Walpole dkk., 2016). Karakteristik dari distribusi *Log-Normal* adalah bernilai positif bukan nol, *skewness* positif dan varians tidak konsisten (heteroskedastisitas) (Gustavsson, 2015).

Variabel acak V berdistribusi *Log-Normal* dinotasikan dengan $V \sim LN(\lambda, \tau^2)$, memiliki fungsi kepadatan peluang pada Persamaan (9).

$$g(v|\lambda,\tau^2) = \frac{1}{\sqrt{2\pi\tau v}} \exp\left(\frac{(\ln(v)-\lambda)^2}{2\tau^2}\right), v > 0\,;\ -\infty \leq \lambda \leq \infty, \tau^2 > 0 \qquad (9)$$

dengan λ merupakan parameter skala dan τ^2 merupakan parameter lokasi. Nilai *mean* dan varians dari distribusi tersebut adalah $E(V) = \exp\left(\lambda + \frac{\tau^2}{2}\right)$ dan $Var(V) = \exp(2\lambda + \tau^2)[\exp(\tau^2) - 1]$.

Jika diberikan variabel random Y yang mengikuti distribusi *Poisson Log-Normal* (μ, τ) maka fungsi kepadatan peluang untuk Y dapat dinyatakan ke dalam bentuk integral pada Persamaan (10) (Ong dkk., 2020).

$$f(y;\mu,\tau) = \int_0^\infty \frac{e^{-\mu v}(\mu v)^y}{y!} g(v)dv, \qquad (10)$$

dengan $g(v)$ adalah fungsi kepadatan peluang bagi $V \sim LN(\lambda, \tau^2)$ pada Persamaan (9).

Apabila fungsi kepadatan peluang variabel acak yang berdistribusi *Log-Normal* pada Persamaan (9) disubsitusikan ke Persamaan (10), maka diperoleh fungsi kepadatan peluang *Poisson Log-Normal* pada Persamaan (11).

$$f(y;\mu,\tau) = \int_0^\infty \frac{e^{-\mu v}(\mu v)^y}{y!} \frac{1}{\sqrt{2\pi\tau v}} \exp\left(-\frac{(\ln(v)-\lambda)^2}{2\tau^2}\right) dv\,; y \geq 0 \qquad (11)$$

Model regresi *Poisson Log-Normal* dengan parameter μ yang dihubungkan dengan variabel prediktor menggunakan fungsi penghubung $\ln(\cdot)$ adalah pada

Persamaan (12) (Perrakis dkk., 2014).

$$\begin{aligned} \ln(\mu_i) &= (\boldsymbol{x}_i^T\boldsymbol{\beta}) \\ \mu_i &= \exp(\boldsymbol{x}_i^T\boldsymbol{\beta}) \end{aligned} \tag{12}$$

dengan:

μ_i : rata-rata variabel respon yang berdistribusi *Poisson*.

$\boldsymbol{x}_i^T$: $[1 \quad x_{i1} \quad x_{i2} \quad \ldots \quad x_{ik}]$ menunjukkan vektor yang berukuran $1 \times (k+1)$ dari variabel prediktor.

$\boldsymbol{\beta}$: $[\beta_0 \quad \beta_1 \quad \beta_2 \quad \ldots \quad \beta_k\,]^T$ menunjukkan vektor yang berukuran $(k+1) \times 1$ dari parameter regresi.

1.5.5. *Bivariate Poisson Log-Normal Regression*

Regresi *Bivarate Poisson Log-Normal* (BPLN) digunakan untuk mengetahui hubungan dua respon berupa data diskrit yang saling berkorelasi dengan beberapa variabel prediktor yang diduga memiliki pengaruh terhadap kedua variabel respon tersebut. Syarat regresi BPLN dapat digunakan adalah kedua variabel respon mengalami overdispersi dan berkorelasi positif. Dua variabel acak (Y_1, Y_2) berdistribusi *Poisson* dan saling bebas, rata-rata $E\left(Y_j\right) = v\mu_j$ dan variansi $Var\left(Y_j\right) = \mu_j + \mu_j^2\tau; j = 1,2$. Jika variabel acak *V* berdistribusi *Log-Normal* Maka fungsi kepadatan peluang bersama Y_1 dan Y_2 ditunjukkan pada Persamaan (13) (Nadarajah & Lyu, 2022).

$$f(y_j; \mu_j, \tau) = \int_0^\infty \prod_{j=1}^2 \frac{\exp(-\mu_j v)\left(\mu_j v\right)^y}{y_j!} g(v)\, dv;\ y_j \geq 0 \tag{13}$$

dengan $g(v)$ adalah fungsi kepadatan peluang bagi $V{\sim}LN(\lambda, \tau^2)$ pada Persamaan (10). Apabila fungsi kepadatan peluang variabel acak yang berdistribusi *Log-Normal* pada Persamaan (9) disubsitusikan ke Persamaan (13), maka diperoleh fungsi kepadatan peluang BPLN pada Persamaan (14).

$$f\left(y_j; \mu_j, \tau; j = 1,2\right) = \int_0^\infty \prod_{j=1}^2 \frac{\exp(-v\mu_j)(v\mu_j)^{y_j}}{y_j!} \frac{1}{\sqrt{2\pi\tau v}} \exp\left(-\frac{\left(\ln(v) + \frac{\tau^2}{2}\right)^2}{2\tau^2}\right) dv \tag{14}$$

Parameter μ_j dihubungkan dengan variabel prediktor menggunakan fungsi penghubung $\ln(\cdot)$. Sehingga model regresi BPLN seperti pada Persamaan (15) (Tzougas & di Cerchiara, 2023).

$$\begin{aligned} \ln(\mu_{ij}) &= (\boldsymbol{x}_i^T\boldsymbol{\beta}_j) \\ \mu_{ij} &= \exp\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_j\right);\ j = 1,2 \end{aligned} \tag{15}$$

dengan:

μ_{ij} : rata-rata variabel respon ke$-j$ yang berdistribusi *Poisson*.

$\boldsymbol{x}_i^T$: $[1 \quad x_{i1} \quad x_{i2} \quad \ldots \quad x_{ik}]$ menunjukkan vektor yang berukuran $1 \times (k+1)$

dari variabel prediktor.

$\boldsymbol{\beta}_j$: $[\beta_{0j} \quad \beta_{1j} \quad \beta_{2j} \quad \dots \quad \beta_{kj}]^T$ menunjukkan vektor yang berukuran $(k+1) \times 1$ dari parameter regresi.

1.5.6. *Maximum Likelihood Estimation*

Maximum Likelihood Estimation (MLE) merupakan metode parameter yang digunakan untuk menaksir parameter suatu model yang diketahui fungsi probabilitasnya. Metode MLE bekerja dengan memanfaatkan fungsi peluang dari parameter yang ditaksir. Misalkan $Y_1, Y_2, \dots, Y_n$ adalah variabel acak dengan fungsi peluang $f(y_i|\boldsymbol{\theta})$ dengan $\boldsymbol{\theta}$ yang merupakan parameter yang tidak diketahui. Karena $Y_1, Y_2, \dots, Y_n$ saling bebas, maka fungsi probabilitas bersama dari $Y_1, Y_2, \dots, Y_n$ dapat dinyatakan sebagai berikut.

$$f(Y_1, Y_2, \dots, Y_n|\boldsymbol{\theta}) = f(y_1|\boldsymbol{\theta}) \cdot f(y_2|\boldsymbol{\theta}) \cdot \dots \cdot f(y_n|\boldsymbol{\theta}) \tag{16}$$

Persamaan (16) dapat dituliskan ke dalam bentuk fungsi *likelihood* yang dinyatakan sebagai (Hogg dkk., 2012):

$$\begin{aligned} L(\boldsymbol{\theta}|y_i) &= f(y_1, y_2, \dots, y_n|\boldsymbol{\theta}) \\ L(\boldsymbol{\theta}|y_i) &= f(y_1|\boldsymbol{\theta}).f(y_2|\boldsymbol{\theta}).\dots.f(y_n|\boldsymbol{\theta}) \\ L(\boldsymbol{\theta}|y_i) &= \textstyle\prod_{i=1}^{n} f(y_i|\boldsymbol{\theta}) \end{aligned} \tag{17}$$

Metode estimasi MLE dilakukan dengan memaksimalkan fungsi *likelihood* $L(\boldsymbol{\theta}|y_i)$. Fungsi *likelihood* dikerjakan dengan logaritma natural yang fungsinya merupakan fungsi monoton naik, sehingga nilai yang memaksimumkan fungsi $l(\boldsymbol{\theta}|y_i)$ sama dengan memaksimumkan fungsi $L(\boldsymbol{\theta}|y_i)$ atau dapat dituliskan seperti pada Persamaan (18).

$$\begin{aligned} l(\boldsymbol{\theta}|y_i) &= \ln L(\boldsymbol{\theta}|y) \\ l(\boldsymbol{\theta}|y_i) &= \ln\{\textstyle\prod_{i=1}^{n} f(y_i|\boldsymbol{\theta})\} \\ l(\boldsymbol{\theta}|y_i) &= \textstyle\sum_{i=1}^{n} \ln f(y_i|\boldsymbol{\theta}) \end{aligned} \tag{18}$$

Nilai parameter $\hat{\boldsymbol{\theta}}$ dapat diperoleh dengan menurunkan $l(\boldsymbol{\theta}|y_i)$ terhadap $\boldsymbol{\theta}$ yang hasilnya disamakan dengan 0 yang ditulis pada Persamaan (19):

$$l'(\boldsymbol{\theta}|y_i) = \frac{\partial l(\boldsymbol{\theta}|y_i)}{\partial \boldsymbol{\theta}} = \mathbf{0} \tag{19}$$

1.5.7. Metode Iterasi *Newton Raphson*

Metode *newton raphson* merupakan metode iterasi untuk menyelesaikan persamaan nonlinier. Metode ini membutuhkan matriks Hessian yang merupakan matriks dengan elemen-elemennya adalah turunan parsial orde dua dari fungsi *likelihood* terhadap masing-masing kombinasi parameter yang digunakan. Persamaan iterasi *newton raphson* untuk parameter $\boldsymbol{\theta}$ dapat dinyatakan melalui Persamaan (20) (Agresti, 2012).

$$\widehat{\boldsymbol{\theta}}^{(s+1)} = \widehat{\boldsymbol{\theta}}^{(s)} - \boldsymbol{H}\left(\boldsymbol{\theta}^{(s)}\right)^{-1} \boldsymbol{g}\left(\boldsymbol{\theta}^{(s)}\right), s = 0,1,2,\dots,q \tag{20}$$

Matriks Hessian parameter $\boldsymbol{\theta}$ pada iterasi ke-s dilambangkan dengan $\boldsymbol{H}(\boldsymbol{\theta}^{(s)})$ pada Persamaan (22).

$$\boldsymbol{g}\left(\boldsymbol{\theta}^{(s)}\right) = \frac{\partial l(\boldsymbol{\theta}|y_i)}{\partial \boldsymbol{\theta}} \tag{21}$$

$$\boldsymbol{H}\left(\boldsymbol{\theta}^{(S)}\right) = \frac{\partial^2 l(\boldsymbol{\theta}|y_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \tag{22}$$

dengan:

$\widehat{\boldsymbol{\theta}}^{(s)}$	:	parameter yang diestimasi pada iterasi ke-s sebanyak q iterasi
$\boldsymbol{g}\left(\boldsymbol{\theta}^{(s)}\right)$	:	Vektor gradien parameter $\boldsymbol{\theta}$ pada iterasi ke-s
$\boldsymbol{H}\left(\boldsymbol{\theta}^{(s)}\right)$	:	Matriks hessian parameter $\boldsymbol{\theta}$ pada iterasi ke-s

Adapun langkah-langkah dalam algoritma *Newton Raphson* adalah sebagai berikut:

1. Menentukan taksiran awal parameter $\widehat{\boldsymbol{\theta}}^{(0)}$ yang diperoleh dari estimasi model regresi *Poisson Log-Normal*.
2. Menghitung vektor gradien $\boldsymbol{g}(\boldsymbol{\theta})$.
3. Menghitung matriks *Hessian* $\boldsymbol{H}(\boldsymbol{\theta})$.
4. Melakukan iterasi menggunakan Persamaan (21) sampai diperoleh nilai yang konvergen, yaitu ketika $\left|\widehat{\boldsymbol{\theta}}^{(s+1)} - \widehat{\boldsymbol{\theta}}^{(s)}\right| \leq \varepsilon$, dengan ε adalah sebesar 0,001.

1.5.8. Pengujian Signifikansi Model *Bivariate Poisson Log-Normal* Regression

1.5.8.1. Uji Simultan

Pengujian signifikansi secara simultan terhadap parameter model regresi BPLN menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) dengan tahapan sebagai berikut.

a. Hipotesis

H_0 : $\beta_{j1} = \beta_{j2} = \cdots = \beta_{jk} = 0, \quad j = 1,2$ dan $\tau = 0$ (variabel prediktor secara bersama-sama tidak mempunyai pengaruh terhadap variabel respon)

H_1 : minimal terdapat $\beta_{jk} \neq 0, j = 1,2\,; k = 1,2,\dots,p$ dan $\tau \neq 0$ (variabel prediktor secara bersama-sama mempunyai pengaruh terhadap variabel respon)

b. Statistik Uji

Pengujian signifikansi parameter secara simultan menggunakan MLRT dirumuskan sebagai berikut:
dengan:

$$G^2_{BPLNR} = -2\ln\left(\frac{L(\widehat{\omega}_{BPLNR})}{L\left(\widehat{\Omega}_{BPLNR}\right)}\right)$$

$$= 2\left(\ln L\left(\widehat{\Omega}_{BPLNR}\right) - \ln L(\widehat{\omega}_{BPLNR})\right) \quad (23)$$

$L(\widehat{\Omega})$: fungsi maksimum *likelihood* untuk model yang melibatkan variabel prediktor.

$L(\widehat{\omega})$: fungsi maksimum *likelihood* untuk model tanpa melibatkan variabel prediktor.

c. Kriteria Uji

Menolak H_0 jika adalah $G^2_{BPLNR} > X^2_{(\alpha,db)}$ dengan db adalah derajat bebas yang diperoleh dari $n(\Omega_{BPLNR}) - n(\omega_{BPLNR})$,artinya variabel prediktor secara bersama sama memiliki pengaruh terhadap variabel respon. (Mardalena dkk., 2021).

1.5.8.2. Uji Parsial

Pengujian secara parsial model regresi BPLN dengan kovariansi sebagai fungsi variabel prediktor digunakan untuk menguji apakah parameter β_{kj} dan τ berpengaruh terhadap model. Pengujian signifikansi dilakukan dengan tahapan sebagai berikut:

a. Hipotesis

H_0 : $\beta_{kj} = 0; k = 1,2,\ldots,p; j = 1,2$ (tidak ada pengaruh secara signifikan)

H_1 : $\beta_{kj} \neq 0$ (Ada pengaruh secara signifikan)

b. Statistik Uji

Pengujian signifikansi parameter secara parsial menggunakan uji *Wald* dirumuskan sebagai berikut:

$$Z_{BPLNR} = \frac{\hat{\beta}_{kj}}{\sqrt{Var(\hat{\beta}_{kj})}} \quad (24)$$

dengan:

$Var(\hat{\beta}_{kj})$: elemen diagonal utama dari matriks varians-kovarians

$Cov(\widehat{\boldsymbol{\theta}}_{\Omega_{BPLNR}})$: matriks varians-kovarians dengan $Cov(\widehat{\boldsymbol{\theta}}_{\Omega_{BPLNR}}) = -\hat{E}\left(\boldsymbol{H}^{-1}\left(\left(\widehat{\boldsymbol{\theta}}_{\Omega_{BPLNR}}\right)\right)\right) = -\boldsymbol{H}^{-1}\left(\left(\widehat{\boldsymbol{\theta}}_{\Omega_{BPLNR}}\right)\right)$ yang bersesuaian dengan $\widehat{\boldsymbol{\theta}}_{\Omega_{BPLNR}} = \left[\widehat{\boldsymbol{\beta}}_{1\Omega_{BPLNR}} \quad \widehat{\boldsymbol{\beta}}_{2\Omega_{BPLNR}} \quad \widehat{\boldsymbol{\tau}}_{\Omega_{BPLNR}}\right]^{\boldsymbol{T}}$.

c. Kriteria Uji

Menolak H_0 jika $|Z_{BPLNR}| > Z_{\alpha/2}$ dengan α adalah taraf signifikan.

1.5.9. Koefisien Korelasi

Koefisien korelasi adalah indikator yang digunakan untuk mengukur hubungan linier antara dua variabel (Draper & Smith, 1992). Koefisien korelasi didefinisikan seperti pada Persamaan (25).

$$r_{y_1,y_2} = \frac{\sum_{i=1}^{n}(y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2)}{\sqrt{\left(\sum_{i=1}^{n}(y_{1i} - \bar{y}_1)^2\right)}\left(\sum_{i=1}^{n}(y_{2i} - \bar{y}_2)^2\right)} \quad (25)$$

Koefisien korelasi dapat menunjukkan dua hubungan, yaitu positif dan negatif yang berada dalam rentang -1 hingga 1. Dua variabel dikatakan memiliki hubungan yang kuat jika nilai korelasinya mendekati 1, baik positif maupun negatif. Sebaliknya, jika nilai korelasinya adalah 0, ini menunjukkan bahwa kedua variabel tidak memiliki hubungan yang signifikan. Korelasi positif menandakan bahwa kedua variabel bergerak searah, sedangkan korelasi negatif menunjukkan bahwa kedua variabel bergerak berlawanan arah (Kawamura, 1973).

Salah satu asumsi yang harus dipenuhi dalam model regresi BPLN adalah pengujian korelasi antar variabel respon. Uji korelasi yang dapat digunakan yaitu uji korelasi *Pearson* dengan hipotesis sebagai berikut:

a. Hipotesis

H_0 : Tidak terdapat hubungan antara Y_1 dan Y_2

H_1 : Terdapat hubungan antara Y_1 dan Y_2

b. Statistik Uji

Pengujian korelasi menggunakan uji korelasi *Pearson* dirumuskan sebagai berikut:

$$t = \frac{r_{y_1,y_2}\sqrt{n-2}}{\sqrt{1-\left(r_{y_1,y_2}\right)^2}} \tag{26}$$

n : banyak pengamatan

c. Kriteria Uji

Menolak H_0 jika $|t_{hit}| > t_{\left(\frac{\alpha}{2};(n-2)\right)}$

1.5.10. Multikolinieritas

Asumsi pada model regresi yang harus dipenuhi adalah tidak terjadi multikolinieritas, yaitu ketika antara variabel prediktor memiliki hubungan korelasi antara yang satu dengan yang lainnya. Variabel prediktor harus saling bebas. Efek terjadinya multikolinieritas pada model adalah nilai *standard erro*r pada model cenderung besar, sehingga meningkatkan peluang untuk gagal menolak H_0 pada pengujian parameter. Multikolinieritas juga menyebabkan matriks variabel prediktor menjadi singular atau tidak *full rank*, sehingga penaksiran parameter tidak dapat dilakukan.

Indikasi terjadi multikolinieritas pada model regresi adalah R^2 tinggi namun tidak ada variabel prediktor yang signifikan. Indikator untuk menentukan terjadinya multikolinieritas adalah dengan *Variance Inflation Factors* (VIF). Pengujian asumsi multikolinieritas dilakukan dengan tahapan sebagai berikut:

a. Hipotesis

H_0 : VIF < 10 (tidak terdapat gejala multikolinieritas pada variabel prediktor).

H_1 : VIF ≥ 10 (terdapat gejala multikolinieritas pada variabel prediktor).

b. Statistik Uji

Nilai VIF untuk variabel prediktor kuantitatif x_k diperoleh dengan Persamaan sebagai berikut.

$$VIF = \frac{1}{1-R_k^2}, k = 1,2,\dots,p \tag{27}$$

dengan:

R_k^2 : koefisien determinasi antara variabel prediktor ke-*k* dengan variabel prediktor lainnya dengan $R_k^2 = \frac{\sum_{i=1}^{n}(\hat{x}_{ki}-\bar{x}_k)^2}{\sum_{i=1}^{n}(x_{ki}-\bar{x}_k)^2}$

x_{ki} : nilai pengamatan ke-*i* untuk variabel prediktor ke-*k*

$\hat{x}_{ki}$: nilai taksiran dari pengamatan ke-*i* untuk variabel prediktor ke-*k*

$\bar{x}_k$: nilai rata-rata variabel prediktor ke-*k*.

c. Kriteria Uji

Menolak H_0 jika $VIF_k \geq 10$. Jika nilai VIF besar maka terjadi multikolinieritas pada data. Pelanggaran asumsi multikolinieritas dapat diatasi dengan menghilangkan variabel yang memiliki kolinieritas tinggi, transformasi data, *ridge regression*, *principal component analysis* (PCA), atau menambah amatan (Gujarati & Porter, 2010).

1.5.11. Kusta

Kusta atau lepra, adalah penyakit infeksi kronis yang disebabkan oleh bakteri *Mycobacterium leprae*. Penyakit ini terutama menyerang kulit, saraf tepi, dan mukosa saluran pernapasan atas yang dalam jangka panjang mengakibatkan sebagian anggota tubuh penderita tidak dapat berfungsi sebagaimana mestinya. Penyakit kusta terdiri dari dua tipe yaitu *Paucl Bacillary* (PB) dan *Multi Bacillary* (MB). Kusta tipe PB adalah tipe kusta yang tidak menular dan disebut juga kusta kering, sedangkan kusta MB atau kusta basah adalah kusta yang sangat mudah menular. Berdasarkan jumlah lesi kulit dan hasil pemeriksaan bakteriologis, kusta PB ditandai dengan adanya lima atau lebih sedikit lesi kulit tanpa bakteri yang terlihat pada pemeriksaan mikroskopis sedangkan Kusta MB ditandai dengan lebih dari lima lesi kulit dan kehadiran bakteri yang dapat terlihat pada pemeriksaan mikroskopis (C. S. Smith dkk., 2017). Tanda-tanda utama penyakit kusta yaitu terjadi kelainan kulit seperti bercak putih, bintik-bintik kemerahan yang tersebar pada kulit, terdapat sebagian dari tubuh yang tidak berkeringat dan mati rasa, serta mengalami perubahan warna kulit menjadi lebih terang atau lebih gelap (Syadifa & Chamid, 2016). Penyakit kusta ditularkan melalui kontak langsung melalui kulit dan saluran pernapasan secara berulang-ulang dan dalam jangka waktu yang lama. Terapi multi-obat yang efektif telah tersedia, dengan deteksi dan pengobatan dini penyakit kusta dapat disembuhkan. Namun penanganan penyakit dapat menjadi rumit karena reaksi yang dimediasi oleh kekebalan tubuh, yang dapat menyebabkan kerusakan saraf permanen dan menyebabkan kecacatan seumur hidup yang berhubungan dengan stigmatisasi dan diskriminasi (Pieter & Grijsen, 2022). Penatalaksanaan kusta yang buruk lainnya dapat mengakibatkan kecacatan, pada mata, tangan, dan kaki (Kemenkes RI, 2022). Faktor risiko penyakit kusta di antaranya adalah kontak serumah dengan penderita kusta, terdapat penderita kusta di lingkungan sekitar tempat tinggal dan kondisi personal hygiene yang buruk (Siswanti & Wijayanti, 2018). Faktor-faktor lain yang diduga menjadi pemicu penyakit kusta adalah kondisi ekonomi keluarga, kebersihan perorangan, lingkungan fisik tempat tinggal dan kepadatan penghuni tempat tinggal (Muharry, 2014).

Berdasarkan laporan dari WHO dalam *Weekly Epidemiological Record* (WER) tahun 2021, Indonesia berada di posisi ketiga sebagai negara dengan jumlah kasus kusta terbanyak di dunia, setelah India dan Brazil. Pada tahun 2021, Indonesia melaporkan 10.876 kasus baru kusta.

Faktor Risiko Penyakit Kusta

Banyak faktor yang memengaruhi terjadinya penyakit kusta bukan saja dari segi medis tetapi meluas sampai masalah sosial, ekonomi dan pendidikan (Dzikirna, 2013).

a. Aspek Kebersihan

Menurut Setiani (2014) perilaku bersih adalah perilaku yang berkaitan dengan upaya mempertahankan dan meningkatkan kesehatan. Penularan penyakit kusta erat kaitannya dengan saluran pernafasan dan kulit. Upaya pencegahan penyakit kusta dapat dilakukan dengan menerapkan rumah tangga berperilaku hidup bersih dan sehat, sehingga dapat meminimalisasi jumlah mikroorganisme penyebab penyakit kusta yang mudah masuk melalui kulit dan saluran pernafasan. Hasil dari penelitian yang telah dilakukan oleh Setiani (2014) adalah seseorang yang memiliki perilaku tidak mencegah penyakit kusta dengan tidak menerapkan upaya pencegahan melalui kebersihan dapat berisiko terkena penyakit kusta sebesar 6,44 kali.

b. Aspek Kepadatan Hunian

Menurut Setiani (2014) seseorang yang memiliki kategori padat penghuni dalam lingkungan rumah dapat berisiko terkena penyakit kusta sebesar 7,22 kali daripada sesorang yang memiliki kategori rumah tidak padat berpenghuni. Hal tersebut terjadi karena kondisi lingkungan rumah yang padat penghuni akan mempermudah penularan penyakit kusta pada orang lain melalui interaksi langsung atau tidak langsung, dimana seorang penderita kusta rata-rata dapat menularkan 2-3 orang di dalam rumahnya. Rumah yang dihuni oleh banyak orang dan ukuran rumah yang tidak sebanding dengan jumlah orang maka akan mengakibatkan dampak buruk bagi kesehatan dan berpotensi terhadap penularan dan infeksi (Rismawati, 2013).

c. Aspek Kemiskinan

Beberapa alasan meningkatnya kejadian kusta pada penduduk miskin adalah penduduk miskin lebih rentan terhadap penyakit karena terbatasnya akses terhadap air bersih dan sanitasi serta kecukupan gizi. Penduduk miskin cenderung enggan mencari pengobatan walaupun sangat membutuhkan karena terdapatnya kesenjangan yang besar dengan petugas kesehatan, terbatasnya sumber daya untuk memenuhi kebutuhan dasar dan terbatasnya pengetahuan untuk menghadapi serangan penyakit (WHO, 2021).

d. Aspek Fasilitas dan Pelayanan Kesehatan

Fasilitas sanitasi rumah perlu ditingkatkan untuk mencegah penyebaran bakteri kusta antara lain pengadaan jamban rumah tangga yang sehat, sarana air bersih yang memenuhi syarat, sarana pembuangan limbah, ventilasi dan pencahayaan yang baik. Abdi (2014) dalam penelitiannya menunjukkan bahwa adanya pengaruh rasio tenaga kesehatan medis per 100.000 penduduk terhadap jumlah penderita kusta.

BAB II
METODE PENELITIAN

2.1. Sumber Data

Data Penelitian ini merupakan studi kasus di Indonesia dengan unit penelitian seluruh Provinsi, sehingga jumlah pengamatan pada penelitian ini adalah 34 provinsi di Indonesia. Penelitian ini menggunakan data sekunder yang diperoleh melalui Laporan Profil Kesehatan Indonesia dari Kementrian Kesehatan Republik Indonesia (Kemenkes RI) yang dapat diakses pada *website* kemkes.go.id dan publikasi Badan Pusat Statistik (BPS) yang dapat di akses pada *website* bps.go.id di Indonesia periode 2021.

2.2. Identifikasi Variabel

Variabel pada penelitian ini terdiri dari dua variabel respon (Y) dan enam variabel prediktor (X) yang diuraikan pada Tabel 1 sebagai berikut.

Tabel 1. Definisi Operasional Variabel Penelitian

Jenis Variabel	Nama Variabel	Definisi Variabel	Satuan
Respon	Kusta PB (Y_1)	Jumlah kasus kusta baru tipe PB tahun 2021.	Kasus
Respon	Kusta MB (Y_2)	Jumlah kasus kusta baru tipe MB tahun 2021.	Kasus
Prediktor	Kepadatan Penduduk (X_1)	Rasio banyaknya penduduk per kilometer persegi.	Jiwa/km^2
Prediktor	Persentase Penduduk Miskin (X_2)	Jumlah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan di bawah garis kemiskinan dibagi dengan jumlah penduduk dikali 100%.	Persen (%)
Prediktor	Persentase Rumah Tangga yang Memiliki Akses Terhadap Layanan Sanitasi Layak (X_3)	Jumlah rumah tangga yang dapat mengakses fasilitas sanitasi layak (fasilitas buang air besar sendiri/bersama, jenis kloset leher angsa, tempat pembuangan berupa septik tank/SPAL) dibagi dengan jumlah rumah tangga keseluruhan dikali 100%.	Persen (%)
Prediktor	Jumlah Fasilitas Kesehatan (X_4)	Jumlah unit fasilitas kesehatan berupa rumah sakit umum, rumah	Unit

Jenis Variabel	Nama Variabel	Definisi Variabel	Satuan
		sakit khusus, puskesmas rawat inap, puskesmas non rawat inap.	
Prediktor	Jumlah Tenaga Medis (X_5)	Jumlah dokter termasuk dokter spesialis (selain spesialis gigi) dan dokter umum.	Individu
Prediktor	Persentase Rumah Tangga yang Memiliki Akses Terhadap Layanan Sumber Air Minum Layak (X_6)	Jumlah rumah tangga yang memiliki akses terhadap air minum layak (air minum dengan jarak ke tempat pembuangan limbah minimal 10 meter yang bersumber dari leding, sumur bor/pompa, sumur terlindung, mata air terlindung, dan termasuk air hujan) dibagi dengan jumlah rumah tangga keseluruhan dikali 100%.	Persen (%)

2.3. Metode Analisis

Tahapan analisis data yang dilakukan berdasarka tujuan penelitian adalah sebagai berikut.

1) Estimasi Parameter Model *Bivariate Poisson Log-Normal Regression*

Tahapan analisis untuk mengestimasi parameter model *Bivariate Poisson Log-Normal Regression* adalah sebagai berikut:

1. Membentuk fungsi *likelihood* untuk pengamatan ke-*i* berdasarkan fungsi massa peluang regresi *Bivariate Poisson Log-Normal* pada Persamaan (15).
2. Menentukan fungsi logaritma natural *likelihood* untuk pengamatan ke-*i* pada Langkah 1.
3. Mencari turunan pertama dari fungsi ln *likelihood* terhadap masing masing parameter yang ditaksir dan disamadengankan nol. Nilai maksimum $L(\boldsymbol{\theta}_{BPLNR})$ akan didapat jika $\frac{\partial l(\boldsymbol{\theta}_{BPLNR})}{\partial \boldsymbol{\theta}_{BPLNR}} = 0$, apabila hasil turunan pertama terhadap masing-masing parameter menghasilkan bentuk implisit maka estimasi parameter diselesaikan menggunakan algoritma *Newton Raphson*.

2) Faktor-Faktor yang Memengaruhi Jumlah Kasus Kusta di Indonesia Menggunakan Model *Bivariate Poisson Log-Normal Regression*

Tahapan analisis untuk membentuk model *Bivariate Poisson Log-Normal Regression* untuk mengetahui faktor-faktor yang memengaruhi jumlah kasus kusta di Indonesia adalah sebagai berikut:

1. Menyajikan statistik deskriptif variabel respon dan variabel prediktor.
2. Menguji distribusi *Bivariate Poisson* pada variabel respon Y_1 dan Y_2

menggunakan *index of dispersion test* pada Persamaan (7).

3. Menguji korelasi antara variabel respon Y_1 dan Y_2 menggunakan uji korelasi *Pearson* (26).
4. Menguji asumsi non-multikolinieritas pada variabel prediktor menggunakan nilai VIF pada Persamaan (27).
5. Melakukan pengecekan overdispersi dengan uji *Deviance* berdasarkan persamaan (8).
6. Mengestimasi parameter model regresi BPLN dengan metode MLE.
7. Melakukan uji signifikansi secara simultan terhadap parameter model BPLNR menggunakan MLRT.
8. Melakukan uji signifikansi secara parsial terhadap parameter $\boldsymbol{\beta}_j$ menggunakan uji *Wald*.
9. Menginterpretasikan model yang diperoleh dan menarik kesimpulan.