REGRESI BINOMIAL NEGATIF PADA KASUS HIV/AIDS DI JAWA TIMUR MENGGUNAKAN ESTIMATOR LOKAL LINIER

SKRIPSI

OLEH
NUCHAILA AINIYAH
NIM. 16610008

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020

REGRESI BINOMIAL NEGATIF PADA KASUS HIV/AIDS DI JAWA TIMUR MENGGUNAKAN ESTIMATOR LOKAL LINIER

SKRIPSI

Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Sains (S.Mat)

Oleh
Nuchaila Ainiyah
NIM. 16610008

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020

REGRESI BINOMIAL NEGATIF PADA KASUS HIV/AIDS DI JAWA TIMUR MENGGUNAKAN ESTIMATOR LOKAL LINIER

SKRIPSI

Oleh
Nuchaila Ainiyah
NIM. 16610008

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal 30 November 2020

Pembimbing I,	Pembimbing II,
Ria Dhea Layla N.K, M.Si	Dr.Hairur Rahman, M.Si
NIDT. 19900709 20180201 2 228	NIP. 19800429 200604 1 003

Mengetahui,
Ketua Jurusan Matematika

Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001

REGRESI BINOMIAL NEGATIF PADA KASUS HIV/AIDS DI JAWA TIMUR MENGGUNAKAN ESTIMATOR LOKAL LINIER

SKRIPSI

Oleh
Nuchaila Ainiyah
NIM. 16610008

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Sains (S.Mat)
Tanggal 30 November 2020

Penguji Utama	: Abdul Aziz, M.Si	....
Ketua Penguji	: Dr. Sri Harini, M.Si	....
Sekretaris Penguji	: Ria Dhea Layla N.K, M.Si	....
Anggota Penguji	: Dr. Hairur Rahman, M.Si	....

Mengetahui,
Ketua Jurusan Matematika

Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama	: Nuchaila Ainiyah
NIM	: 16610008
Jurusan	: Matematika
Fakultas	: Sains dan Teknologi
Judul Skripsi	: Regresi Binomial Negatif Pada Kasus HIV/AIDS di Jawa Timur Menggunakan Estimator Lokal Linier

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 30 November 2020
Yang membuat pernyataan,

METERAI TEMPEL
6000
ENAM RIBU RUPIAH

Nuchaila Ainiyah
NIM. 16610008

MOTTO

Jika kamu benar-benar menginginkan sesuatu, lambat launkamu pasti akan segera menemukan caranya.

PERSEMBAHAN

Skripsi ini penulis persembahkan untuk:

Ibunda Khoirul Abidah, Ayahanda Choidin, serta adik tersayang Achmad Muzammi Fahmi. Semua keluarga besar yang telah mendukung setiap langkah penulis dan yang selalu memberikan semangat.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji bagi Allah SWT atas rahmat, taufik serta hidayah-Nya, sehingga penulis mampu menyelesaikan penyusunan skripsi ini sebagai salah satu syarat untuk memperoleh gelar sarjana dalam bidang matematika di Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Dalam proses penyusunan skripsi ini, penulis banyak mendapat bimbingan dan arahan dari berbagai pihak. Untuk itu ucapan terima kasih yang sebesar-besarnya penulis sampaikan terutama kepada:

1. Prof. Dr. H. Abd Haris, M.Ag, selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Usman Pagalay, M.Si, selaku ketua Jurusan Matematika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Ria Dhea Layla N. K., M.Si, selaku dosen pembimbing I yang telah banyak memberikan arahan, nasihat, motivasi dan berbagi ilmunya kepada penulis.
5. Dr. Hairur Rahman, M.Si, selaku dosen pembimbing II yang telah memberikan arahan, nasihat, motivasi dan berbagi ilmunya kepada penulis.
6. Segenap civitas akademika Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
7. Bapak dan Ibu serta adik tercinta yang selalu memberikan do'a, semangat, serta motivasi kepada penulis sampai saat ini.
8. Semua pihak yang tidak dapat disebutkan satu persatu, yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa masih terdapat kekurangan yang harus disempurnakan dari skripsi ini. Oleh karena itu, penuulis mengharap kritikan dan masukan yang dapat membangun dan meningkatkan kualitas skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi kepentingan ilmu di masa depan.

Malang, 30 November 2020

Penulis

DAFTAR ISI

BAB III METODOLOGI PENELITIAN

BAB IV HASIL DAN PEMBAHASAN

BAB V PENUTUP

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR SIMBOL

y_i	: Data variabel respon ke-i
x_i	: Data variabel prediktor ke-i
e	: *Exponensial*
α	: Parameter *disperse*
β	: Nilai konstanta rill
H_0	: Hipotesis awal
H_1	: Hipotesis *alternative*
α	: Taraf signifikan
LR	: Uji *Likelihood Ratio*
K_{hj}	: Kernel pembobot
$L(\beta, \alpha)$	: Fungsi *Likelihood*

ABSTRAK

Nuchaila, Ainiyah. 2020. **Regresi Binomial Negatif Pada Kasus HIV/AIDS di Jawa Timur Menggunakan Estimator Lokal Linier.** Tugas akhir/skripsi. Jurusan Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Ria Dhe Layla N.K, M.Si (II) Dr. Hairur Rahman, M.Si.

Kata Kunci: Overdispersion, Binomial Negatif, HIV/AIDS, MLCV, Bandwidth

Overdispersion terjadi dalam suatu model poisson jika variansi dari variable respon lebih besar dari mean. Regresi binomial negatif dapat digunkan untuk mengatasi kasus overdispersion. Model regresi binomial negatif dengan pendekatan regresi nonparametric berdasarkan estimator lokal linier. Metode yang digunakan untuk penaksiran regresi $m(x_i)$ pada titik x_0 adalah dengan metode *Locally Weighted Maximum Likelihood Estimator*. fungsi *local likelihood* terboboti tidak dapat diselesaikan secara langsung, maka digunkan iterasi *Newton Rapshon*. Metode untuk mendapatkan *h* optimal dapat dilakukan menggunakan metode *Maximum Likelihood Cross Validation* (MLCV). Tujuan dari penelitian ini adalah mencari estimasi parameter serta model regresi binomial negatif menggunakan estimator lokal linier . data yang digunakan adalah banyaknya kasus HIV/AIDS di Jawa timur pada tahun 2018 dengan faktor yang mempengaruhi adalah banyak penduduk yang mendonorkan darah dan banyaknya pengguna NAPZA. Model regresi yang diperoleh adalah $\hat{y} = \exp(6{,}29602 + 0{,}000987x_1 + 0{,}002765x_2)$ didapat bandwidth optimal sebesar 200 dan 8 dengan nilai MLCV(h) maksimum sebesar 29,5410 dan 52,8617.

ABSTRACT

Nuchaila, Ainiyah. 2020. Negative Binomial Regression in HIV/AIDS Cases in East Java Using a Local Linear Estimator. Thesis. Departement of Mathematics, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University of Malang, Advisors: (I) Ria Dhea Layla N.K, M.Si (II) Dr. Hairur Rahman, M.Si.

Keywords: Overdispersion, Negative Binomial, HIV/AIDS, MLCV, Bandwidth

Overdispersion occurs in a Poisson model if the variance of the response variable is greater than the mean. Negative binomial regression can be used to treat overdispersion cases. Negative binomial regression model with nonparametric regression approach based on linear local estimator. The method used to estimate point regression is themethod *Locally Weighted Maximum Likelihood Estimator*. Thefunction *local likelihood* weightedcannot be solved directly, so we use theiteration *Newton Rapshon*. The method to obtain *h* optimalcan be done using themethod *Maximum Likelihood Cross Validation* (MLCV). The purpose of this study is to find parameter estimates and negative binomial regression models using a linear local estimator. The data used is the number of HIV / AIDS cases in East Java in 2018 with factors affecting the number of people donating blood and the number of drug users. The regression model obtained is $\hat{y} = \exp(6{,}29602 + 0{,}000987x_1 + 0{,}002765x_2)$ obtained the optimal bandwidth is 200 and 8 with the maximum MLCV(h) values of 29.5410 and 52,8617.

الملخّص

نحيلة, عينية,2020. **الانحسار الرمزي السلبي في قصيّة فيروس نقص المناعة (أي الإيدز) بجاوي الشرقي باستخدام المقدار الداخليّ الموفقي**. كلية العلوم و التكنولوجيا. جامعة مولانا مالك إبراهيم الاسلامية الحكومية مالانج. المشرفة (١) ريبا ديبا نور حاريسما الماجستير, المشرف (٢) الدكتور خير الرحمن الماجستير.

الكلمات المرشدة: الانتشار المفرط, الرمزي السلبي , فيروس نقص المناعة (أي الإيدز), MLCV , Bandwidth.

وقع الانتشار المفرط في الشكل فويسون إذا كان النوع و الاجابة المقلّب أكبر من " مين ". و الانحسار الرمزي السلبي مستخدم لتسديد مشكلة الانتشار المفرط. الانحسار الرمزي السلبي بطريقة الانحسار غير المعامل بالمقدار الداخليّ الموفقي. و الطريقة المستخدمة لتقدير الانحسار $M(x_1)$ في نقطة X_0 هي بالطريقة المقدار الثقلي الداخليّ و مهمتها لا تتمّ مباشرة و لذلك استخدم الوصف الحاسوبي نوطان و رافسون. و الطريقة لنيل " h "الكاملة هي بالطريقة التحقيقية المتفوقة (MLCV). أمّا الهدف من هذا البحث هو بحث تقدير المعامل و شكل الانحسار الرمزي السلبي باستخدام المقدار الداخليّ الموفقي. و الحقائق المستخدمة هي كثرة مشكلة فيروس نقص المناعة (أي الإيدز) في جاوى الشرقي سنة 2018 و العوامل التي تسبّبها هي السكاّن الذين أعطوا دماءهم كثير منهم الذين يتناولون المخدرات أي NAPZA . و تنال شكل الانكسار " $\hat{y} = \exp(6{,}29602 + 0{,}000987x1 + 0{,}002765x2\ )$ " و تنال " Bandwidth " الكاملة على قدر 200 و 8 بنتيجة الطريقة التحقيقية المتفوقة / " MLCV(h) " على قدر التفويق 29,5410 و 52,8617 .

BAB I
PENDAHULUAN

1.1 Latar Belakang

Kesehatan adalah hak dasar manusia dan merupakan salah satu aspek kualitas sumberdaya manusia yang penting untuk dicermati. Hal itu dikarenakan kesehatan merupakan salah satu peran penting dalam investasi sumber daya manusia, maka sangat tepat peran yang dilakukan pemerintah dalam pembangunan di bidan kesehatan secara terus menerus. *Human Immunodeficiency Virus* (HIV)/*Acquired Immunodeficiency Syndrome* (AIDS) HIV/AIDS telah menjadi masalah internasional, karena dalam waktu yang relatif singkat terjadi peningkatan jumlah penderita dan semakin melanda banyak negara. Perkembangan HIV/AIDS di Indonesia dari tahun ke tahun terus meningkat.

Penyebaran HIV/AIDS ini terus meningkat sangat dipengaruhi oleh faktor seperti penggunaan alat kontrasepsi, pendonoran darah, penggunaan NAPZA dan jumlah tenaga medis disetiap daerah. Permasalahan HIV/AIDS menjadi tantangan kesehatan hampir diseluruh dunia, termasuk di Indonesia. Sejak pertama kali ditemukan sampai dengan Juni 2018, HIV/AIDS telah dilaorkan keberadaannya oleh 433 (84,2%) dari 514 kabupaten/kota di 34 provinsi di Indonesia. Adapun provinsi dengan jumlah infeksi HIV tertinggi adalah DKI Jakarta (55.099), diikuti Jawa Timur (43.399), Jawa Barat (31.293), Papua (30.699) dan Jawa Tengah (24.757) (Kemenkes, 2018).

Dalam pandangan Islam, sakit merupakan musibah yang dapat menimpa siapa saja, termasuk orang-orang yang shaleh dan berakhlak mulia sekalipun. Artinya, orang yang terkena penyakit belum tentu sakitnya itu akibat perbuatan dosa yang dilakukan, tetapi boleh jadi merupakan korban perbuatan orang lain. Pada dasarnya Islam sarat dengan tuntunan untuk berpola hidup sehat secara jasmani dan rohani. Diantaranya, Islam mengajarkan untuk menghindari penyakit dan berobat jika sakit, bersabar dan banyak beristighfar jika mendapat musibah, merawat serta pemperlakukan orang yang sakit dengan baik.

Apabila sedang tertimpa musibah, termasuk jika sedang sakit, kita diperintahkan untuk banyak bersabar sambil bertobat/berikhtiar. Allah SWT berfirman :

... وَاصْبِرْعَلَي مَا اَصَابَكَ, إِنَّ ذَلِكَ مِنْ عَزْمِ الْأُمُوْرَ

Artinya: *"... dan bersabarlah atas apa yang menimpa kamu, sesungguhnya yang demikian itu termasuk hal-hal yang mewajibkan (oleh Allah)." (QS Luqman:17).*

Meski demikian, kita juga harus menghindari hal-hal yang dapat membahayakan diri sendiri atau orang lain, termasuk juga harus hati-hati terhapat penyakit yang menular. Salah satu penyakit yang menular yaitu HIV/AIDS, penyebabnya adalah berzina dalam pegertian yang luas yang menurut ajaran Islam merupakan perbuatan keji yang diharamkan oleh Allah SWT.

وَلاَتَقْرَبُواْ الزِّني صلي إِنَّهُ كَانَ فَحِشَةً وَسَآءَ سَبِيْلاً

Artinya: *"Dan janganlah kamu mendekati zina, sesungguhnya zina itu adalah suatu perbuatan yang keji dan suatu jalan yang buruk." (QS. Al-Isra 17:32).*

Ayat diatas mengisyaratkan bahwa Islam melarang segala jenis kegiatan yang mengarah kepada perzinaan. Mengingat HIV/AIDS sebagian besar diakibatkan oleh perzinaan termasuk diantaranya seks pra nikah, prostitusi, homoseks dan penggunaan narkoba.

Estimasi model yang digunakan untuk memodelkan pada penyakit HIV/AIDS adalah binomial negatif untuk mengatasi kasus *overdispersion* dengan regresi nonparametrik dan melibatkan variabel prediktor berdasarkan estimator lokal linier. Pada regresi Poisson terdapat asumsi yang harus dipenuhi, yaitu varians dan rata-rata dari variabel respon tersebut sama atau *equidispersion*. Namun dalam kenyataannya di lapangan sering terjadi pelanggaran asumsi tersebut, yaitu nilai variansnya lebih besar dari nilai rata-rata yang dinamakan overdispersion atau nilai variansnya lebih kecil dari nilai rata-rata yang dinamakan underdispersion (Wang dan Famoye, 1997). Hal ini jelas melanggar asumsi pada regresi Poisson dan apabila regresi Poisson tetap digunakan, maka dapat mengakibatkan kesalahan dalam inferensi bagi parameter model karena penduga nilai variansi akan lebih kecil dari nilai variansi yang sebenarnya. (Astuti dan Yanagawa, 2002).

Berbagai penelitian untuk mengatasi kasus overdispersion pada model regresi Poisson telah banyak dilakukan. Penelitian tersebut Greene (2008) dan Masson (2014) mengembangkan regresi binomial negatif untuk mengatasi kasus *overdispersion*. Penelitian lain dilakukan oleh Ismail dan Jemain (2007) membahas penggunaan *Generalized Poisson* dan *Negative Binomial Regression* yang keduanya terbukti dapat mengatasi kasus overdispersion pada model regresi Poisson. Penelitian lain oleh Asrul dan Naing (2012) menganalisis perbandingan regresi binomial negatif dan Poisson pada data tingkat kematian AIDS dan diperoleh hasil bahwa regresi binomial negatif lebih baik daripada regresi Poisson dalam mengatasi kasus *overdispersion*.

Model regresi pada data *count* untuk mengatasi kasus *overdispersion* yang telah disebutkan diatas merupakan regresi parametrik yang mengasumsikan fungsi regresinya mengikuti bentuk atau pola tertentu. Faktanya hubungan antara variabel respon, khususnya berupa data *count*, dengan variabel prediktor tidak selalu menunjukkan pola tertentu pada bentuk kurva regresinya. Pada saat kondisi seperti ini, model regresi nonparametrik lebih cocok untuk digunakan (Eubank, 1999). Bentuk kurva atau fungsi regresi nonparametrik diasumsikan mulus atau smooth, sehingga pendekatan pada regresi nonparametrik memberikan fleksibilitas yang lebih besar dibanding regresi parametrik karena bentuk estimasi fungsi regresi akan menyesuaikan dengan pola data tanpa dipengaruhi subyektifitas peneliti (Hardle, 2004).

Pengembangan model regresi nonparametrik dengan variabel respon berupa data *count* telah dilakuan oleh Santos dan Neves (2008) yang membahas tentang regresi Poisson dengan estimator lokal linier dan menggunkan metode *local maximum likelihood* dalam pengestimasian parameternya. Adapun dalam hal mengatasi kasus *overdispersion* terdapat penelitian oleh Bary dan Welsh (2002) yang melakukan pemodelan jumlah dahan pohon *Eucalyptus mannifera* dengan beberapa variabel prediktor menggunkan regresi binomial negatif terpotong dan GAM (*Generalized Additif Model*). Peneliti lain oleh Aulia (2014) yang membahas penerapan regresi poisson dan binomial negatif dalam memodelkan jumlah kasus penderita AIDS di Indonesia berdasarkan faktor sosiodemografi. Penelitian lain oleh Astuti (2013) yang mengembangkan model regresi nonparametrik untuk data count yakni dengan menggunakan *local generalized Poisson regression* berdasarkan estimator polinomial lokal untuk memodelkan angka kematian Indonesian dengan 1 variabel prediktor.

Regresi nonparametrik merupakan teknik analisis data yang dapat digunkan untuk menjelaskan hubungan fungsional variabel pengamatan dan variabel prediktor yang tidak mengasumsikan bentuk tertentu dari kurva regresi. Bentuk kurva atau fungsi regresi nonparametrik diasumsikan mulus atau *smooth*, sehingga regresi nonparametrik memberikan fleksibilitas yang tinggi dalam menentukan bentuk fungsi regresinya (Eubank, 1998). Beberapa estimator utuk mengestimasi kurva atau fungsi regresi nonparametrik, antara lain adalah estimator kernel, estimator lokal linier, estimator polinomial lokal, estmator *spline truncated*, estimator *spline penalized* dan estimator *deret fourier*. Estimator lokal linier dipilih karena mempunyai kelebihan dapat mengestimasi fungsi disetiap titik sehingga hasil estimasi yang didapat lebih mendekati pola data sebenarnya. Selain itu, estimator lokal linier juga memenuhi prinsip parsimoni sehingga data dapat dinterpretasi dengan mudah (Nottingham dan Cook, 2001).

Berdasarkan beberapa hal yang telah dipaparkan diatas, dapat disimpulkan bahwa belum ada penelitian mengenai estimasi model regresi binomial negatif nonparametrik berdasarkan estimator lokal linier, peneliti juga akan mengimplementasikan teori tersebut pada data real bidang kesehatan. Dinas kesehatan, sebagai pihak yang menjalankan kebijakan dan program pengembangan kesehatan, harus berorientasi pada target *Milenium Development Goals* (MDG's). Salah satu tujuan pokok MDG's adalah memerangi HIV/AIDS. HIV/AIDS adalah penyakit menular seksual (PMS) yang tidak dapat disembuhkan dan paling berbahaya karena dapat merusak kekebalan tubuh seseorang dan dapat menyebabkan kematian.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka rumusan masalah dalam penelitian skripsi ini adalah:

1. Bagaimana estimasi regresi binomial negatif menggunakan estimator lokal linier?
2. Bagaimana modela regresi binomial negatif pada kasus HIV/AIDS di Jawa Timur menggunakan estimator lokal linier?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dikemukakan sebelumya maka tujuan penelitian skripsi ini adalah:

1. Mendapatkan estimasi regresi binomial negatif menggunakan estimator lokal linier.
2. Mendapatkan model regresi binomial negatif pada kasus HIV/AIDS di Jawa Timur menggunakan estimator lokal linier.

1.4 Batasan Masalah

Berdasarkan rumusan masalah dan tujuan penelitian, batasan masalah dalam penelitian ini adalah:

Penelitian ini diaplikasikan pada jumlah kasus HIV/AIDS di Jawa Timur dengan variabel yang digunakan adalah banyaknya penderita HIV/AIDS di Jawa Timur (Y), banyaknya pengguna alat kontrasepsi (X_1) dan banyaknya pendonor darah (X_2).

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat, khususnya kepada penulis dan umumnya kepada semua pembaca baik secara teoritis maupun secara praktis.

1. Menambah pengetahuan dan wawasan tentang regresi binomial negatif menggunakan estimator lokal linier.
2. Menambah pengetahuan tentang regresi binomial negatif pada kasus HIV/AIDS di Jawa Timur menggunakan estimator lokal linier.

1.6 Sistematika Penulisan

Dalam penulisan skripsi ini, penulis menggunakan sistematika penulisan yang terdiri dari lima bab dan masing – masing bab memiliki beberapa subbab dengan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan mengenai pendahuluan yang didalamnya terdapat latar belakang yang melatar belakangi penulis memilih judul, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini menjelaskan mengenai hal-hal yang mendasari teori yang digunakan peneliti dan memaparkan pustaka yang dipakai pada saat penelitian dan penulisan. Teori yang diambil dari jurnal, literatur, komunikasi dan internet.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan mengenai data dan sumber data, variabel penelitian, langkah-langkah penelitian dan flowchart

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan hasil dan pembahasan pada penelitian.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi uraian tentang pokok-pokok kesimpulan dan saran-saran yang perlu disampaikan kepada pihak yang berkepentingan.

BAB II
TINJAUAN PUSTAKA

2.1 *Overdispersion* dan Distribusi Binomial Negatif

Overdispersion terjadi dalam suatu model Poisson jika variansi dari variabel respon lebih besar dari Mean. *Overdispersion* tersebut dipengaruhi oleh korelasi yang positif diantara variabel respon atau dari variansi yang berlebihan pada probabilitas variabel respon. *Overdispersion* menjadi suatu masalah karena dapat menyebabkan standard error dari variabel dugaan menjadi *underestimated* yaitu variabel prediktor yang faktanya tidak signifikan dapat menjadi signifikan dan dalam hal tersebut menjadi sebuah kesalahan dalam penarikan kesimpulan. Salah satu upaya untuk mengatasi kasus *Overdispersion* pada data adalah dengan menggunakan distribusi Binomial Negatif (Hilbe, 2011).

Distribusi binomial negatif merupakan distribusi yang bisa didekati dengan barisan Bernoulli dan distribusi Poisson-Gamma. Distribusi Binomial Negatif yang dibentuk dari Poisson-Gamma memiliki fungsi sebagai berikut:

$$f(y_i,\mu,\alpha) = \begin{cases} \dfrac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)}\left(\dfrac{1}{1+\alpha\mu}\right)^{\frac{1}{\alpha}}\left(\dfrac{\alpha\mu}{1+\alpha\mu}\right)^{y_i} & ,i = 0,1,2,\ldots \\ 0\,, & \text{untuk selainnya} \end{cases} \quad (2.1)$$

dengan μ adalah parameter lokasi yang bernilai lebih dari 0 dan α adalah parameter dispersi (Boswell dan Patil, 1970).

Hilbe (2011) menyatakan bahwa distribusi dari x merupakan anggota dari kelompok eksponensial jika fungsi probabilitasnya memiliki bentuk sebagai berikut:

$$f(y,\theta,\phi) = \exp\left\{\frac{x\theta - b(\theta)}{a(\phi)} + c(y;\phi)\right\} \quad (2.2)$$

dengan θ merupakan parameter kanonik, $a(\phi)$ merupakan parameter skala dan $\mathrm{b}(\theta)$ adalah fungsi limit cumulant.

Salah satu keuntungan dari kelompok eksponensial adalah mudah untuk mengidentifikasi mean dan variansi. Mean dari distribusi yang merupakan anggota dari kelompok eksponensial adalah turunan pertama dari $\mathrm{b}(\theta)$ dan variansinya adalah hasil turunan kedua dari $\mathrm{b}(\theta)$ (Hilbe, 2011).

Distribusi Binomial Negatif merupakan kelompok eksponensial. Bukti akan diberikan dengan menyesuaikan persamaan (2.1) dengan persamaan (2.2) seperti berikut:

$$f(y_i,\mu,\alpha)=\left(\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)}\right)\left(\frac{1}{1+\alpha\mu}\right)^{\frac{1}{\alpha}}\left(\frac{\alpha\mu}{1+\alpha\mu}\right)^{y_t}$$

$$=\exp\left\{\ln\left(\left(\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)}\right)\left(\frac{1}{1+\alpha\mu}\right)^{\frac{1}{\alpha}}\left(\frac{\alpha\mu}{1+\alpha\mu}\right)^{y_t}\right)\right\}$$

$$=\exp\left\{\ln\left(\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)}\right)+\frac{1}{\alpha}\ln\left(\frac{1}{1+\alpha\mu}\right)+y_i\ln\left(\frac{\alpha\mu}{1+\alpha\mu}\right)\right\}$$

$$f(y_i,\mu,\alpha)=\exp\left\{\frac{y_i\ln\left(\frac{\alpha\mu}{1+\alpha\mu}\right)+\frac{1}{\alpha}\ln\left(\frac{1}{1+\alpha\mu}\right)}{1}+\ln\left(\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)}\right)\right\} \quad (2.3)$$

Pada persamaan di atas, diperoleh $\theta=\ln\left(\frac{\alpha\mu}{1+\alpha\mu}\right), \alpha(\phi)=1, b(\theta)=-\frac{1}{\alpha}\ln\left(\frac{1}{1+\alpha\mu}\right)$, dan $c(y;\phi)=\ln\left(\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i!\Gamma\left(\frac{1}{\alpha}\right)}\right)$

Berdasarkan Hilbe (2010), pengujian parameter dispersi dapat dilakukan dengan menggunakan uji *Likelihood Ratio* dengan hipotesis sebagai berikut:

H_0 : $\alpha=0$ (tidak terjadi *overdispersion*)

H_1 : $\alpha\neq 0$ (terjadi *overdispersion*)

Statistik uji yang digunakan adalah *Likelihood Ratio* dengan persamaan sebagai berikut:

$$LR=-2(\ell_p-\ell_{NB}) \quad (2.4)$$

Kriteria penolakan H_0 adalah jika nilai $LR>x^2_{(\alpha,p)}$ dengan p adalah banyak prediktor dan α adalah tingkat signifikasi yang digunakan.

2.2 Regresi Binomial Negatif

Regresi Binomial Negatif merupakan salah satu pendekatan untuk mengatasi variabel respon data diskrit yang nilai variansinya lebih besar daripada mean atau disebut *Overdispersion*. Model dari regresi binomial negatif dapat dilihat pada persamaan berikut:

$$f(y_i|x_i) = \left(\frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)}\right)\left(\frac{1}{1+\alpha\mu(x_i)}\right)^{\frac{1}{\alpha}}\left(\frac{\alpha\mu(x_i)}{1+\alpha\mu(x_i)}\right)^{y_i} \quad (2.5)$$

dengan $y_i = 0,1,2,..,\alpha \geq 0$ dan $\mu(x_i)$ didefinisikan sebagai persamaan berikut:

$$\mu(x_i) = \mathrm{E}(y_i|x_i) = m(x_i) \quad (2.6)$$

Dalam regresi binomial negatif, y_i adalah variabel respon data diskrit yang bernilai positif, maka nilai ekspektasi dari y_i juga bernilai positiff. Hal tersebut menjadi sesuatu yang menyimpang karena ruang nilai μ_i pada persamaan (2.5) berada pada interval $(-\infty, \infty)$.

Untuk mengatasi penyimpangan tersebut, maka digunakan sebuah fungsi yang menghubungkan antara $\mu(x_i)$ dan $m(x_i)$. Hilbe (2011) menyatakan bahwa model binomial negatif pada umumnya menggunakan fungsi penghubung logaritma atau log link. Oleh sebab itu, dengan ditambah fungsi logaritma , maka persamaan (2.6) dapat ditulis menjadi:

$$\ln\big(\mu(x_i)\big) = m(x_i) \quad (2.7)$$

Persamaan diatas menjadi tepat karena akan terdefinisi didalam interval $(0, \infty)$ dan interpretasi menjadi mudah. Setelah diperoleh fungsi penghubung yang tepat, berdasarkan persamaan (2.7) maka diperoleh persamaan $\mu(x_i)$ sebagai berikut:

$$\mu(x_i) = \exp\big(m(x_i)\big) \quad (2.8)$$

Dengan $m(x_i)$ adalah fungsi regresi pada regresi parametrik yang didefinisikan sebagai $\boldsymbol{x_i\beta}$.

2.3 Regresi Nonparametrik

Regresi nonparametrik merupakan suatu metode statistika yang digunakan untuk mengetahui hubungan antara variabel respon dan variabel prediktor ketika bentuk kurva regresi antara variabel respon dan variabel prediktor tidak diketahui atau tidak didapatkan informasi sebelumnya. Hubungan antara variabel respon y dan variabel prediktor x dapat dinyatakan dalam persamaan berikut:

$$y_i = m(x_i) + \varepsilon_i, i = 1,2, \ldots n, \varepsilon_i - (0, \sigma^2) \quad (2.9)$$

Dengan $m(x_i)$ adalah fungsi regresi yang akan diestimasi dan diasumsikan smooth sehingga lebih menjamin fleksibilitas dalam mengestimasi fungsi regresinya (Eubank, 1999).

Sebagai perluasan dari persamaan (2.9), hubungan antara variabel respon y dan variabel multiprediktor adalah sebagai berikut:

$$y_i = \sum_{j=1}^{p} m(x_{ij}) + \varepsilon_i, \qquad i = 1,2,\ldots,n, \varepsilon_i \sim (0,\sigma^2) \tag{2.10}$$

2.4 Fungsi Kernel

Salah satu pembobotan yang digunakan untuk mendapatkan estimasi adalah fungsi kernel (Eubank, 1998). Kelebihan dari estimator kernel adalah memiliki bentuk lebih fleksibel dan secara matematik mudah disesuaikan serta mempunyai rata-rata kekonvergenan yang relatif cepat (Hardle, 1990). Fungsi kernel K dengan *bandwidth h* didefinisikan sebagai berikut:

$$K_h(t) = \left(\frac{1}{h}\right) K\left(\frac{t}{h}\right); -\infty < t < \infty \text{ dan } h < 0 \tag{2.11}$$

Sifat-sifat dari fungsi kernel adalah sebagai berikut:

1. $K(t) \geq 0$ untuk semua t
2. $\int_{-\infty}^{\infty} K(t)dt = 1$
3. $\int_{-\infty}^{\infty} tK(t)dt = 0$
4. $\int_{-\infty}^{\infty} t^2K(t)dt = \sigma^2 > 0$

Sehingga estimator fungsi densitas kernelnya adalah:

$$\hat{f}_t(t) = \frac{1}{n}\sum_{i=1}^{n} K_h(t - t_i) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{t - t_1}{h}\right) \tag{2.12}$$

Estimator fungsi densitas kernel pada persamaan (2.12) mempunyai tujuh macam fungsi kernel yang digunakan. Fungsi kernel tersebut dapat dilihat pada tabel 2.1:

Tabel 2. 1 Jenis Fungsi Kernel

No	Nama Fungsi	Bentuk Fungsi

1	Kernel Uniform	$K(t) = \frac{1}{2}I(\lvert t\rvert \leq 1)$
2	Kernel Segitiga	$K(t) = (1 - \lvert t\rvert)I(\lvert t\rvert \leq 1)$
3	Kernel Epanechnikov	$K(t) = \frac{3}{4}(1 - t^2)I(\lvert t\rvert \leq 1)$
4	Kernel Kuadrat	$K(t) = \frac{15}{16}(1 - r^2)I(\lvert t\rvert \leq 1)$
5	Kernel Triweight	$K(t) = \frac{35}{32}(1 - r^2)^3 I(\lvert t\rvert \leq 1)$
6	Kernel Cosinus	$K(t) = \frac{\pi}{4}\cos\left(\frac{\pi}{2}t\right) I(\lvert t\rvert \leq 1)$
7	Kernel Gaussian	$K(t) = \frac{1}{\sqrt{2\pi}} exp\left(\frac{1}{2}(-t^2)\right), -\infty < t < \infty$

2.5 Estimator Lokal Linier

Berdasarkan Fan dan Gijbels (1996), estimator lokal linier merupakan salah satu pendekatan regresi nonparametrik yang menggunakan teknik smoothing. Estimator lokal linier mempunyai kelebihan dibandingkan estimator yang lain yaitu dapat mengestimasi fungsi di setiap titik sehingga model yang didapatkan lebih mendekati pola data yang sesungguhnya dan estimator ini tidak membutuhkan data dalam jumlah yang banyak dalam pengestimasian modelnya (Nottingham dan Cook, 2001).

Fungsi regresi $m(x_{ij})$ pada persamaan (2.9) merupakan fungsi yang tidak diketahui atau tidak terikat pada asumsi fungsi tertentu dan diestimasi dengan pendekatan non parametrik menggunakan estimator lokal linier. Fungsi regresi $m(x_{ij})$ yang dihasilkan merupakan fungsi *smooth* yang memiliki sifat kontinu dan diferensiabel. Oleh karena itu, untuk mengestimasinya dapat didekati dengan ekspansi *taylor* pada x disekitar titik x_0, yakni:

$$m(x_i) \approx \sum_{k=0}^{d} \frac{m^{(k)}(x_0)}{k!}(x_i - x_0)^k, x_i \in (x_0 - h, x_0 + h) \tag{2.13}$$

Persamaan di atas merupakan persamaan polinomial berorde d. Oleh karena lokal linier merupakan kasus khusus polinomial lokal dengan orde 1, maka berdasarkan persamaan (2.13), maka fungsi regresi yang akan diestimasi menggunakan estimator lokal linier dijabarkan pada persamaan berikut:

$$m(x_i) \approx \frac{m^{(0)}(x_0)}{0!}(x_i - x_0)^0 + \frac{m^{(1)}(x_0)}{1!}(x_i - x_0)^1, x_i \in (x_0 - h, x_0 + h)$$

$$\approx \frac{m^{(0)}(x_0)}{0!} + \frac{m^{(1)}(x_0)}{1!}(x_i - x_0)^1, x_i \in (x_0 - h, x_0 + h) \qquad (2.14)$$

Jika didefinisikan $\frac{m^{(k)}(x_0)}{k!} = \beta_k(x_0)$, maka persamaan (2.14) dapat ditulis menjadi persamaan berikut:

$$m(x_i) = \beta_0(x_0) + \beta_1(x_0)(x_i - x_0), x_i \in (x_0 - h, x_0 + h) \qquad (2.15)$$

Jika dinyatakan dalam notasi matriks menjadi:

$$\boldsymbol{m(x_i) = x\beta} \qquad (2.16)$$

dengan $x = \left(1(x_i - x_0)\right)$ dan $\beta = \begin{pmatrix} \beta_0(x_0) \\ \beta_0(x_0) \end{pmatrix}$

2.6 Metode *Locally Weighted Maximum Likelihood Estimator*

Dalam regresi nonparametrik, salah satu metode yang dapat digunakan untuk penaksiran kurva regresi $m(x_i)$ pada titik x_0 dengan terlebih dahulu mengestimasi parameternya adalah dengan metodel *Locally Weighted Maximum Likelihood Estimator*. Menurut Astuti (2013), fungsi *local likelihood* terboboti dalam regresi nonparametrik berdasarkan estimator lokal linier mempunyai persamaan sebagai berikut:

$$L(\beta, \alpha, x_0) = \prod_{i=1}^{n} f\,(y_i|x_i)^{k_h(x_i - x_0)} \qquad (2.17)$$

Sedangkan bentuk fungsi local likelihood terboboti untuk kasus variabel multiprediktor sebagai perluasan dari persamaan (2.17) dapat ditulis pada persamaan berikut:

$$L(\beta, \alpha, x_0) = \prod_{i=1}^{n} \left\{ f(y_i|X)^{\prod_{j=1}^{p} K_{ij}(x_{ij} - x_{0j})} \right\} \qquad (2.18)$$

Oleh karena nilai $L(\beta, \alpha, x_0)$ dari persamaan (2.18) diperoleh fungsi log-likelihood sebagai berikut:

$$\ln\{\mathrm{L}(\beta, \alpha, x_0)\} = \ln\left\{ \prod_{i=1}^{n} \left\{ f(y_i|X)^{\prod_{j=1}^{p} K_{ij}(x_{ij} - x_{0j})} \right\} \right\} \qquad (2.19)$$

Estimator $\hat{\beta}$ diperoleh dari turunan pertama fungsi persamaan (2.19) terhadap β dan pada estimator $\hat{\alpha}$ terhadap α kemudian disama dengankan 0. Sedemikian sehingga mendapatkan persamaan sebagai berikut:

$$\frac{\partial \ell(\beta, \alpha, x_0)}{\partial \beta} = 0 \tag{2.20}$$

$$\frac{\partial \ell(\beta, \alpha, x_0)}{\partial \alpha} = 0 \tag{2.21}$$

2.7 Metode *Newton Raphson*

Menurut Park dan Lord (2008), untuk menaksir parameter dari regresi binomial negatif digunakan metode maksimum likelihood dengan prosedur literasi Newton Raphson. Metode Newton Raphson adalah salah satu metode iteratif yang digunakan untuk menyelesaikan persamaan yang tidak dapat diselesaikan secara langsung karena tidak linier dalam parameternya (Cameron dan Trivedi, 2008).

Berdasarkan Agresti (2002), langkah-langkah literasi menggunakan metode newton raphson untuk mengestimasi parameter adalah sebagai berikut:

a. Menentuka nilai estimasi awal parameter yakni $\hat{\beta}^*_{(0)} = \begin{bmatrix} \hat{\beta}_{0(0)} \\ \hat{\beta}_{1(0)} \\ \vdots \\ \hat{\beta}_{p(0)} \\ \hat{\alpha}_{(0)} \end{bmatrix}$

b. Membentuk vektor gradient g sesuai pada persamaan berikut:

$$g^T \left(\frac{\partial \ell(\beta, \alpha, x_0)}{\partial \beta_0}, \frac{\partial \ell(\beta, \alpha, x_0)}{\partial \beta_0}, \dots, \frac{\partial \ell(\beta, \alpha, x_0)}{\partial \beta_p}, \frac{\partial \ell(\beta, \alpha, x_0)}{\partial \alpha} \right) \tag{2.21}$$

c. Membentuk matriks hessian H, sesuai persamaan berikut:

$$H = \begin{pmatrix} \frac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0^2} & \cdots & \frac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0 \partial \beta_p} & \frac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0 \partial \alpha} \\ & \ddots & \vdots & \frac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_p \partial \alpha} \\ & & \frac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0^2} & \frac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \alpha^2} \end{pmatrix} \tag{2.22}$$

Untuk dapat memaksimumkan fungsi likelihood, matriks H harus merupakan matriks definit negatif.

d. Memasukkan nilai $\hat{\beta}^*_{(0)}$ kedalam elemen-elemen vektor gradient g dan matriks hessian H.

e. Menghitung invers dari matriks hessian (H^{-1}).

f. Melakukan iterasi m dimulai dari 0 pada persamaan:

$$\beta_{(m-1)} = \beta_{(m)} - H^{-1}, m = 0,1,2, \dots \quad (2.23)$$

g. Jika belum didapat estimator parameter yang konvergen, iterasi kembali dilakukan hingga m=m+1. Iterasi berhenti jika nilai $|\beta_{(m-1)} - \beta_{(m)}| \leq \varepsilon$.

2.8 Turunan Fungsi Log Gamma

Fungsi gamma merupakan suatu fungsi khusus. Fungsi ini dapat digunakan untuk menyederhanakan integral-integral khusus. Fungsi gamma yang dinotasikan sebagai $\Gamma(n)$ didefinisikan sebagai:

$$\Gamma(n) = \int_0^\infty x^{n-1}\, e^{-x} dx, n > 0 \quad (2.24)$$

Berdasarkan Hilbe (2011), fungsi di gamma adalah turunan pertama dari fungsi log gamma yang didefinisikan sebagai berikut:

$$\begin{aligned}\psi(n) &= \frac{d \ln \Gamma(n)}{d(n)} \\ &= \frac{\Gamma'(n)}{\Gamma(n)} \\ &= \frac{\ln \Gamma(n + 0{,}0001) - \ln \Gamma(n - 0{,}0001)}{0{,}0002} \quad (2.25)\end{aligned}$$

Sedangkan fungsi trigamma adalah turunan kedua dari fungsi log gamma yang didefinisikan sebagai berikut:

$$\begin{aligned}\psi'^{(n)} &= \frac{d^2 \ln \Gamma(n)}{d(n)^2} \\ &= \frac{\Gamma''^{(n)}}{\Gamma(n)} \\ &= \frac{-\ln \Gamma(n + 0{,}002) + 16^* \ln \Gamma(n + 0{,}001)}{0{,}000012} \\ &+ \frac{-30 \ln \Gamma(n) + 16 \ln \Gamma(n - 0{,}001) - \ln \Gamma\,(n - 0{,}002)}{0{,}000012} \quad (2.26)\end{aligned}$$

2.9 *Maximum Likelihood Cross Validation* (MLCV)

Bandwidth (h) adalah parameter penghalus yang digunakan untuk mengontrol kehalusan kurva yang diestimasi. Apabila h yang dipilih terlalu kecil maka akan menghasilkan kurva yang sangat kasar dan fluaktif. Sebaliknya, apabila h yang dipilih terlalu lebar akan menghasilkan kurva yang sangat mulus dan tidak dapat merepresentasikan data. Oleh karena itu, memilih nilai h optimal merupakan hal yang penting (Liang dan Cheng, 2005).

Pada model yang menggunakan metode maximum likelihood untuk mendapatkan estimasi kurva regresinya, maka metode untuk mendapatkan h optimal dapat dilakukan menggunakan metode Maximum Likelihood Cross Validation (MLCV). MLCV yang merupakan fungsi dari h yang dijabarkan dalam persamaan berikut:

$$MLCV(h) = \sum_{i=1}^{n} \ln f\left(y_i, m_{-i}(x)\right) \tag{2.27}$$

Dengan $m_{-i}(x)$ merupakan estimasi fungsi regresi pada titik x_i tanpa menyertakan data ke-i.

2.10 Statistik Uji *Deviance*

Statistik uji *deviance* merupakan salah satu uji statistik yang dapat digunakan untuk menguji kesesuaian model (goodness of fit) pada regresi binomial negatif. Secara umum, rumus statistik uji Deviance didefinisikan sebagai berikut:

$$D(y, \hat{\mu}) = 2\{\ell(y) - \ell(\hat{\mu})\} \tag{2.28}$$

dengan $\ell(y)$ adalah fungsi log likelihood dengan setiap nilai dan μ adalah nilai dari y itu sendiri, sedangkan $\ell(\bar{\mu})$ adalah fungsi log likelihood untuk model yang diestimasi.

Untuk menguji kesesuain model regresi binomial negatif dengan statistik uji Deviance, hipotesis yang digunakan adalah sebagai berikut:

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

H_1 : minimal ada satu $\neq 0$

Berdasarkan Cameron dan Trivedi (1998), statistik uji deviance untuk model regresi binomialnegatif dengan parameter dispersi (α) adalah sebagai berikut:

$$\mathrm{D}_{NB} = \sum_{i=1}^{n}\left\{y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i + \alpha^{-1})\ln\left[\frac{y_i + \alpha^{-1}}{\hat{\mu}_i + \alpha^{-1}}\right]\right\} \qquad (2.29)$$

Adapun kriteria uji yang digunakan adalah H_0 ditolak jika $D_{NB} > \chi^2_{(n-p,\alpha)}$, dengan n adalah banyak observasi, p adalah banyak prediktor dan α adalah tingkat signifikansi yang digunakan (Famoye, 2010).

2.11 HIV/AIDS dan Faktor yang Berpengaruh

Acquired Immunodeficiency Syndrome (AIDS) adalah infeksi yang disebabkan oleh *Human Immunodeficiency Virus* (HIV) yang menyebabkan suatu penyakit yang menyerang sel-sel kekebalan tubuh. Virus HIV/AIDS masuk ke dalam tubuh manusia melalu perantara darah, semen dan sekret vagina. HIV/AIDS tergolong retnovirus yang mempunyai materi genetika RNA yang mampu menginfeksi limfosit CD4 (*Cluster Differential Four*), dengan melakukan perubahan sesuai dengan DNA inangnya (Kementrian Kesehatan, 2014).

Berdasarkan hasil statistik kasus HIV/AIDS yang dilaporkan oleh Ditjen Pengendalian Penyakit (PP) dan Penyehatan Lingkungan (PL), jumlah kasus HIV/AIDS mengalami peningkatan terdapat 35 juta penduduk seluruh dunia hidup dengan HIV. Indonesia dikategorikan sebagai negara dengan tingkat penular HIV/AIDS tertinggi. Jawa timur merupakan salah satu dari 5 provinsi dengan kasus HIV/AIDS kumulatif terbanyak di Indonesia (Dinkes Jatim, 2017).

Banyaknya kasus HIV/AIDS tentunya tidak terlepas dari berbagai faktor yang mempengaruhinya. Beberapa faktor yang mempengaruhi kasus HIV/AIDS di Jawa Timur dijelaskan sebagai berikut:

1. Banyaknya Penduduk yang Mendonorkan Darah

Donor darah merupakan kegiatan penyaluran darah dari satu orang ke sistem peredaran orang lain. Kegiatan donor darah tidak bebas dari resiko. Hal tersebut karena meskipun semua donor darah telah diperiksa, namun resiko penularan bahan penularan seperti HIV/AIDS dan virus hepatitis tidak dapat dipastika sepenuhnya bahwa tidak akan terjadi (O'brien, 2006). Dinas Kesehatan Provinsi Jawa Timur (2015) menyatakan jumlah keseluruhan baik laki-laki maupun perempuan sebesar 455.536. seluruh donor darah tersebut diambil sampel darah yang diperiksa terhadap HIV/AIDS. Hasil pemeriksaan menunjukkan bahwa terdapat 1.025 donor darah yang positif mengandung HIV/AIDS.

2. Penggunaan NAPZA

Narkotika adalah zat atu obat yang berasal dari tanaman atau bukan tanaman, baik sintesis maupun semi sintesis, yang dapat menyebabkan penurunan atau erubahan kesadaran, hilangnya rasa, mengurangi sampai menghilangkan rasa nyeri dan dapat menimbulkan ketergantungan. Narkotika dan obat terlarang serta zat adiktif psikotropika dapat menyebabkan efek dan dan dampak negatif bagi pemakainya. Menurut Kemenkes RI (2017), penggunaan narkotika dapat menyebabkan penyakit menular berbahaya seperti HIV/AIDS. Ditjen P2P Kemenkes RI(2017) menyatakan pada kasus yang dilaporkan tahn 2016, proporsi penderita HIV disebabkan oleh pengguna narkoba adalah sebesar 1.9 persen dari banyak kasus HIV/AIDS yang ada. Menurut kepala BNN Jawa Timur, penggunaan narkoba di Jawa Timur mencapai 0,02 persen dari seluruh jumlah penduduk di Jawa Timur total 40 juta jiwa.

BAB III
METODOLOGI PENELITIAN

3.1 Pendekatan Penelitian

Pendekatan penelitian yang digunakan dalam penelitian ini adalah dengan pendekatan studi literatur dan deskriptif kuantitatif. Pada studi literatur menggunakan bahan-bahan pustaka yang dibutuhkan penulis sebagai acuan untuk menyelesaikan penelitian. Deskriptif kuantitatif menggunakan data sekunder yang sesuai dengan kebutuhan peneliti.

3.2 Sumber Data

Data yang digunakan adalah data yang diperoleh dari Dinas Kesehatan Provinsi Jawa Timur tahun 2018. Variabel yang digunakan adalah 38 Kabupaten/Kota di Provinsi Jawa Timur. Data diambil pada hari Kamis, 20 Februari 2020 pada pukul 19.00 WIB.

3.3 Identifikasi Variabel

Variabel penelitian yang digunakan pada penelitian ini adalah sebagai berikut:

Tabel 3. 1 Variabel Penelitian

No	Variabel	Keterangan	Skala Data
1	y	Banyaknya Kasus HIV/AIDS di tiap Kabupaten dan Kota di Jawa Timur	Diskrit
2	x_1	Banyaknya penduduk yang mendonorkan darah di tiap Kabupaten dan Kota di Jawa Timur	Diskrit
3	x_2	Banyaknya pengguna NAPZA di tiap Kabupaten dan Kota di Jawa Timur	Diskrit

3.4 Metode Analisis Data

3.4.1 Estimasi Model Regresi Binomial Negatif Menggunakan Estimator Lokal Linier

Langkah-langkah yang dilakukan untuk mengestimasi model sebagai berikut:

1. Mengidentifikasi data yang berpasangan (x_i, y_i) dimana x_i adalah variabel prediktor pada data observasi ke-i dan y_i adalah variabel respon pada data observasi ke-i dengan asumsi distribusi binomial negatif berdasarkan pada persamaan (2.1).
2. Menyatakan model distribusi binomial negatif multiprediktor $f(y_i|x_i)$ yang sesuai pada persamaan (2.5).
3. Mensubsitusikan $\mu(x_i)$ pada langkah yang ke 2 sesuai dengan persamaan yang merupakan perluasan dari $\mu(x_i)$ pada persamaan (2.6).
4. Mendeteksi fungsi regresi $m(x_i)$ pada langkah ke 3 dengan pendekatan lokal linier multiprediktor $m(x_{ij})$ yang merupakan perluasa dari $m(x_i)$ pada persamaan 2.14
5. Menyatakan hasil pada langkah 4 pada bentuk matriks
6. Mensubsitusikan notasi matriks hasil dari langkah ke 5 kedalam persamaan yang didapatkan pada langkah ke 3.
7. Menentukan fungsi *log-likelihood* distribusi binomial negatif dengan pendekatan lokal linier prediktor berdasarkan persamaan (2.17) dengan $f(y_i|x_i)$ pada persamaan yang didapat dari langkah ke 6.
8. Menentukan turunan pertama terhadap β pada persamaan (2.20) dan turunan pertama pada α sesuai persamaan (2.21) dengan mensubsitusi $\ell(\beta, \alpha)$ pada persamaan hasil langkah ke 7
9. Menentukan turunan kedua terhadap α dan β

3.4.2 Pemodelan Regresi Binomial Negatif pada Data HIV/AIDS di Jawa Timur Menggunakan Estimator Lokal Linier

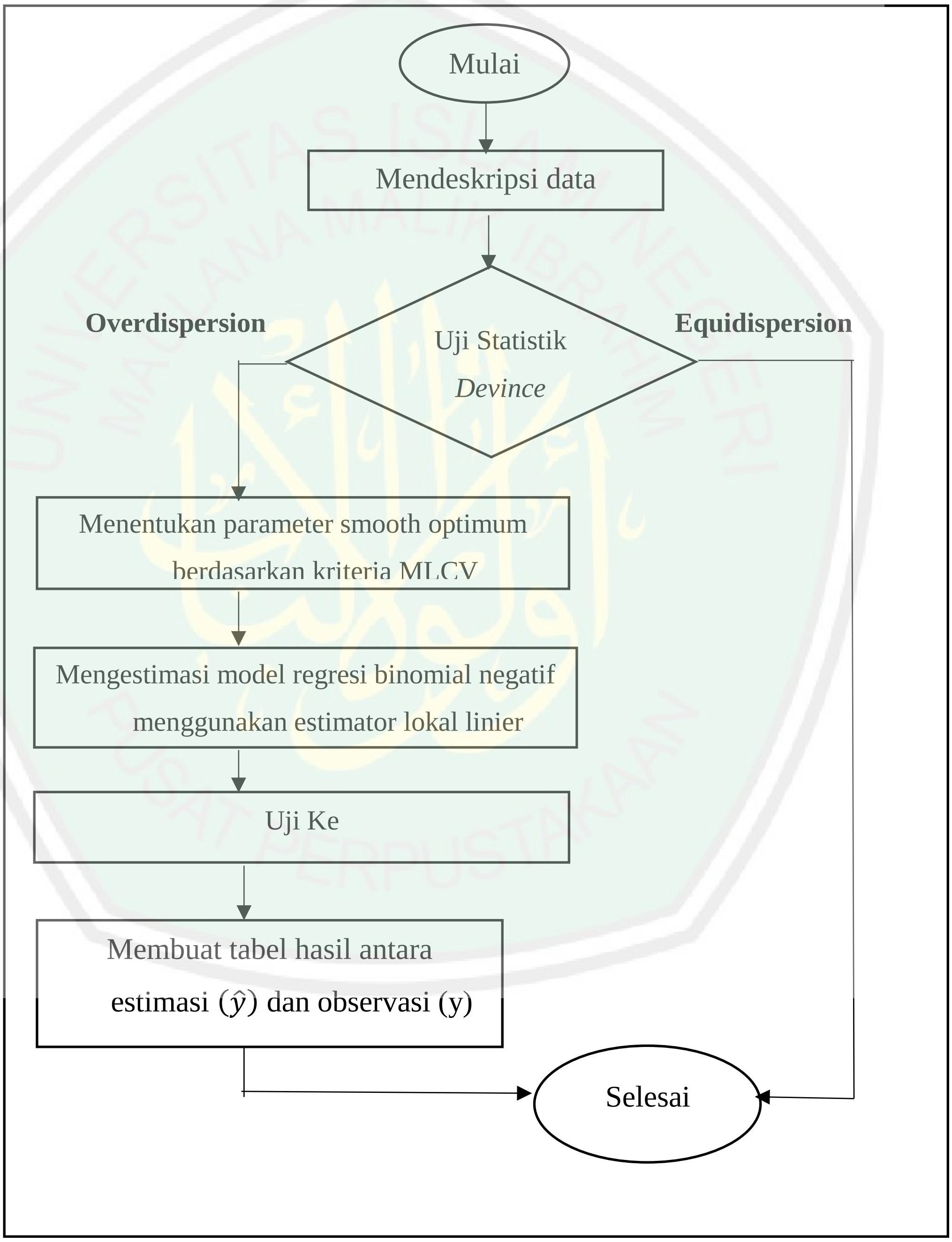
Langkah-langkah untuk mengestimasikan model regresi nonparametrik binomial negatif menggunakan estimator lokal linier adalah sebagai berikut:

1. Menginput data berpasangan (y_i, x_i).
2. Mengidentifikasi terjadinya *overdispersion* terhadap variabel respon.
3. Menghitung nilai MLCV (h) untuk mendapatkan bandwith optimal.
4. Mengestimasi fungsi regresi berdasarkan metode regresi nonparametrik binomial negatif multiprediktor menggunakan estimator lokal linier dengan parameter *smooth* yang diperoleh pada langkah ke 3.

5. Membuat tabel pada hasil estimasi ($\hat{y}$)dan observasi (y)
6. Menghitung nilai *deviance* untuk menguji kesesuaian model.
7. Melakukan analisis dan interpretasi terhadap model yang dihasilkan

3.5 Flow Chart

Berikut Flow Chart atau diagram alir pada penelitian ini



Gambar 3. 1 Diagram Alir Model Regresi Binomial Negatif Nonparametrik berdasarkan Estimator Lokal Linier

BAB IV

HASIL DAN PEMBAHASAN

4.1 Estimasi Model Regresi Binomial Negatif Nonparametrik berdasarkan Estimator Lokal Linier

Diketahui pasangan data (x_i, y_i) dengan x_i menyatakan variabel predictor pada observasi ke-i dan y_i menyatakan variabel respon pada observasi ke-i. Data berpasangan terseut merupakan variabel random tipe diskrit yang diasumsikan berdistribusi binomial negatif dan mempunyai model regresi pada persamaan (2.1).

Pada persamaan (4.1), diketahui bahwa $\mu(x_i)$ bergantung pada variabel x_i melalui suatu fungsi persamaan (2.7) sebagai berikut:

$$\mu(x_i) = \exp\big(\mathrm{m}(x_i)\big) \tag{4.1}$$

Melalui subsitusi persamaan (4.2) dan (4.1), maka persamaan (4.1) berubah menjadi:

$$f(x_i|y_i) = \left(\frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)}\right)\left(\frac{1}{1 + \alpha \exp\big(\mathrm{m}(x_i)\big)}\right)^{\frac{1}{\alpha}}\left(\frac{\alpha \exp\big(\mathrm{m}(x_i)\big)}{1 + \alpha \exp\big(\mathrm{m}(x_i)\big)}\right)^{y_i} \tag{4.2}$$

Pada persamaan di atas, $\mathrm{m}(\mathrm{x_i})$ merupakan fungsi yang tidak diketahui atau tidak terikat pada asumsi bentuk fungsi tertentu. Fungsi tersebut didekati regresi nonparametrik berdasarkan estimator lokal linier yakni didefinisikan sebagai berikut:

$$\sum_{j=1}^{p} m(x_{ij}) = \sum_{j=1}^{p}\{\beta_{0j}(x_{0j}) + \beta_{1j}(x_{0j})(x_{ij} - x_{0j})\}, x_{ij} \in (x_{0j} - h, x_{0j} + h) \tag{4.3}$$

Dengan penjabaran sebagai berikut:

$$\sum_{j=1}^{p} m(x_{1j}) = \beta_{01}(x_{01}) + \beta_{11}(x_{01})(x_{11} - x_{01}) + \beta_{02}(x_{02}) + \beta_{12}(x_{02})(x_{12} - x_{02}) + \cdots + \beta_{0p}(x_{0p}) + \beta_{1p}(x_{0p})(x_{1p} - x_{0p})$$

$$\sum_{j=1}^{p} m(x_{2j}) = \beta_{01}(x_{01}) + \beta_{11}(x_{01})(x_{21} - x_{01}) + \beta_{02}(x_{02}) + \beta_{12}(x_{02})(x_{22} - x_{02}) + \cdots + \beta_{0p}(x_{0p}) + \beta_{1p}(x_{0p})(x_{2p} - x_{0p})$$

$$\vdots$$

$$\sum_{j=1}^{p} m(x_{nj}) = \beta_{01}(x_{01}) + \beta_{11}(x_{01})(x_{n1} - x_{01}) + \beta_{02}(x_{02}) + \beta_{12}(x_{02})(x_{n2} - x_{02}) + \cdots + \beta_{0p}(x_{0p}) + \beta_{1p}(x_{0p})(x_{np} - x_{0p})$$

apabila dinyatakan dalam bentuk matriks, didapatkan persamaan berikut:

$$\mathrm{m}(\mathrm{x_i}) = \mathrm{X}\beta \qquad (4.4)$$

dengan

$$\mathrm{m}(\mathrm{x_i}) = \begin{pmatrix} \sum_{j=1}^{p} m(x_{1j}) \\ \sum_{j=1}^{p} m(x_{2j}) \\ \vdots \\ \sum_{j=1}^{p} m(x_{nj}) \end{pmatrix}$$

$$\mathbf{X} = \begin{pmatrix} \mathrm{x_1} \\ \mathrm{x_2} \\ \mathrm{x_3} \\ \vdots \\ \mathrm{x}_p \end{pmatrix}$$

$$= \begin{pmatrix} (1 & (x_{11} - x_{01}) & (x_{12} - x_{02}) & \cdots & (x_{1p} - x_{0p})) \\ (1 & (x_{21} - x_{01}) & (x_{22} - x_{02}) & \cdots & (x_{2p} - x_{0p})) \\ & & \vdots & & \\ (1 & (x_{n1} - x_{01}) & (x_{n2} - x_{02}) & \cdots & (x_{np} - x_{0p})) \end{pmatrix} \text{ dan}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \sum_{j=1}^{p} \beta_0(x_{0j}) \\ \beta_{11}(x_{01}) \\ \beta_{12}(x_{02}) \\ \vdots \\ \beta_{1p}(x_{0p}) \end{pmatrix}$$

Langkah selanjutnya dilakukan dengan mensubsitusikan persamaan (4.4) pada persamaan (4.3) diperoleh persamaan sebagai berikut:

$$f(y_i|x_i) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{1}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha \exp\,(\mathrm{x}_i\beta)}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right)^{y_i} \tag{4.5}$$

Dengan tidak mengubah bentuk yang ada, maka persamaan (4.5) dapat ditulis sebagai berikut:

$$f(y_i|x_i) = \exp \ln \left\{ \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{1}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha \exp\,(\mathrm{x}_i\beta)}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right)^{y_i} \right\}$$

$$= \exp\left\{ \ln\left(\frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)} \right) + \frac{1}{\alpha} \ln\left(\frac{1}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right) + y_i \ln\left(\frac{\alpha \exp\,(\mathrm{x}_i\beta)}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right) \right\}$$

$$= \exp\left\{ \ln \Gamma\left(y_i + \frac{1}{\alpha}\right) \ln y_i!\,\Gamma\left(\frac{1}{\alpha}\right) - \frac{1}{\alpha}\ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) + y_i \ln\left(\alpha \exp\,(\mathrm{x}_i\beta)\right) - y_i \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) \right\}$$

$$= \exp\left\{ \ln \Gamma\left(y_i + \frac{1}{\alpha}\right) - \ln y_i!\,\Gamma\left(\frac{1}{\alpha}\right) - \frac{1}{\alpha}\ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) + y_i \ln \alpha + y_i(\mathrm{x}_i\beta) - y_i \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) \right\} \tag{4.6}$$

Untuk mengestimasi parameter β dalam penelitian ini digunakan metode Maksimum *Local Likelihood* terboboti. Adapun fungsi local likelihood terboboti dengan fungsi kernel multiprediktor dapat diberikan pada persamaan yang merupakan perluasan dari persamaan (2.19) pada kasus multiprediktor sebagai berikut:

$$\ell(\beta, \alpha) = \ln\left\{ \prod_{i=1}^{n} \left\{ f(y_i|x_i)^{\prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j})} \right\} \right\} \tag{4.7}$$

pada persamaan (4.7), komponen $f(y_i|x_i)$ disubsitusikan dengan persamaan (4.6) sehingga persamaan (4.7) ditulis menjadi:

$$\ell(\beta, \alpha) = \ln\left\{ \prod_{i=1}^{n} \left\{ \exp\left\{ \ln \Gamma\left(y_i + \frac{1}{\alpha}\right) - \ln y_i! - \Gamma\left(\frac{1}{\alpha}\right) - \frac{1}{\alpha}\ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) + y_i \ln \alpha + y_i(\mathrm{x}_i\beta) - y_i \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)^{\prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j})} \right\} \right\} \right\}$$

$$= \ln\left\{\exp \sum_{i=1}^{n} \left\{\ln \Gamma\left(y_i + \frac{1}{\alpha}\right) - \ln y_i - ! \, \Gamma\left(\frac{1}{\alpha}\right) - \frac{1}{\alpha} \ln (1 + \alpha \exp(\mathrm{x}_i\beta)) + y_i \ln \alpha + y_i(\mathrm{x}_i\beta) - y_i \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)\right\}^{\prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j})}\right\}$$

$$= \ln\left\{\exp \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left\{\ln \Gamma\left(y_i + \frac{1}{\alpha}\right) - \ln y_i! \, \Gamma\left(\frac{1}{\alpha}\right) - \frac{1}{\alpha} \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) + y_i \ln \alpha + y_i(\mathrm{x}_i\beta) - y_i \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)\right\}\right\}$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left\{\ln \Gamma\left(y_i + \frac{1}{\alpha}\right) - \ln y_i! - \ln \Gamma\left(\frac{1}{\alpha}\right) - \frac{1}{\alpha} \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) + y_i \ln \alpha + y_i(\mathrm{x}_i\beta) - y_i \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)\right\} \tag{4.8}$$

Dalam pengestimasian parameter β, fungsi *local likelihood* pada persamaan (4.8) kemudian diturunkan terhadap β dengan penjabaran sebagai berikut:

$$\frac{\partial \ell(\beta, \alpha,)}{\partial \beta} = \left\{\sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left(0 - 0 - 0 - \frac{1}{\alpha}\left(\frac{\alpha \exp(\mathrm{x}_i\beta)}{1 + \exp(\mathrm{x}_i\beta)} x_i + 0 + y_i(x_i) - + \frac{y_i \alpha \exp(\mathrm{x}_i\beta).(\mathrm{x}_i)}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right)\right)\right\}$$

$$= \left\{\sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left(\frac{-\exp(\mathrm{x}_i\beta).(\mathrm{x}_i)}{1 + \alpha \exp(\mathrm{x}_i\beta)} + \frac{y_i(x_i)\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right) + \frac{-y_i.(x_i).\alpha \exp(\mathrm{x}_i\beta)}{1 + \alpha \exp(\mathrm{x}_i\beta)}\right\}$$

$$= \left\{ \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left(\frac{-\exp(\mathrm{x}_i\beta).(\mathrm{x}_i) + y_i(x_i)}{1 + \alpha \exp\,(\mathrm{x}_i\beta)} + \frac{y_i.(\mathrm{x}_i).\alpha \exp(\mathrm{x}_i\beta)}{1 + \alpha \exp\,(\mathrm{x}_i\beta)} + \frac{-y_i.(x_i).\alpha \exp(\mathrm{x}_i\beta)}{1 + \alpha \exp\,(\mathrm{x}_i\beta)} \right) \right\}$$

$$= \left\{ \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left(\frac{-\exp(\mathrm{x}_i\beta)(\mathrm{x}_i) + y_i(x_i)}{1 + \alpha \exp\,(\mathrm{x}_i\beta)} \right) \right\} \tag{4.9}$$

Sedangkan untuk mengestimasi parameter dispersi (α), dilakukan penurunan fungsi *local likelihood* terhadap α, yang dijabarkan sebagai berikut:

$$\frac{\partial \ell(\beta, \alpha)}{\partial \alpha} = \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left\{ \psi\left(y_i + \frac{1}{\alpha}\right) - 0 - \psi\left(\frac{1}{\alpha}\right) - \left(-\frac{1}{\alpha^2} \ln(1 + \alpha \exp(\mathrm{x}_i\beta)) + \frac{1}{\alpha}\left(\frac{\exp(\mathrm{x}_i\beta)}{1 + \alpha \exp(\mathrm{x}_i\beta)} \right) \right) + \frac{y_i}{\alpha} + 0 - \frac{y_i(\exp(\mathrm{x}_i\beta))}{1 + \alpha \exp(\mathrm{x}_i\beta)} \right\}$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left\{ \psi\left(y_i + \frac{1}{\alpha}\right) - \psi\left(\frac{1}{\alpha}\right) + \frac{y_i}{\alpha} - \frac{y_i(\exp(\mathrm{x}_i\beta))}{1 + \alpha \exp(\mathrm{x}_i\beta)} + \frac{1}{\alpha^2} \ln(1 + \alpha \exp(\mathrm{x}_i\beta)) - \frac{1}{\alpha} \frac{\exp(\mathrm{x}_i\beta)}{(1 + \alpha \exp(\mathrm{x}_i\beta))} \right\}$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left\{ \psi\left(y_i + \frac{1}{\alpha}\right) - \psi\left(\frac{1}{\alpha}\right) + \frac{y_i\alpha(1 + \alpha \exp(\mathrm{x}_i\beta))}{\alpha^2(1 + \alpha \exp(\mathrm{x}_i\beta))} + \frac{y_i\alpha^2(\exp(\mathrm{x}_i\beta))}{\alpha^2(1 + \alpha \exp(\mathrm{x}_i\beta))} - \frac{(1 + \alpha \exp(\mathrm{x}_i\beta))\ln(1 + \alpha \exp(\mathrm{x}_i\beta))}{\alpha^2(1 + \alpha \exp(\mathrm{x}_i\beta))} - \frac{\alpha(\exp(\mathrm{x}_i\beta))}{\alpha^2(1 + \alpha \exp(\mathrm{x}_i\beta))} \right\}$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left\{ \psi\left(y_i + \frac{1}{\alpha}\right) - \psi\left(\frac{1}{\alpha}\right) + \frac{y_i\alpha + y_i\alpha^2\left(\exp(\mathrm{x}_i\beta)\right)}{\alpha^2\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)} \right.$$

$$- \frac{y_i\alpha^2\left(\exp(\mathrm{x}_i\beta)\right)}{\alpha^2\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)} + \frac{1}{\alpha^2}\ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)$$

$$\left. - \frac{\alpha\left(\exp(\mathrm{x}_i\beta)\right)}{\alpha^2\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)} \right\}$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left\{ \psi\left(y_i + \frac{1}{\alpha}\right) - \psi\left(\frac{1}{\alpha}\right) + \frac{y_i\alpha - \alpha\left(\exp(\mathrm{x}_i\beta)\right)}{\alpha^2\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)} \right.$$

$$\left. + \frac{1}{\alpha^2}\ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) \right\}$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left\{ \psi\left(y_i + \frac{1}{\alpha}\right) - \psi\left(\frac{1}{\alpha}\right) \right.$$

$$\left. + \frac{1}{\alpha^2}\left(\frac{\alpha\{y_i - \left(\exp(\mathrm{x}_i\beta)\right)\}}{\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)} + \ln\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right) \right) \right\} \quad (4.10)$$

Nilai maksimum dari fungsi *local likelihood* pada persamaan (4.8) akan tercapai saat persamaan (4.9) dan persamaan (4.10) sama , namun karena kedua persamaan tersebut tidak dapat diselesaikan secara langsung, maka diperlukan suatu metode pendekatan untuk mendapatkan penyelesaiannya. Salah satu metode numerik yang dapat digunakan adalah metode *Newton Raphson*.

Dalam metode Newton Raphson, beberapa komponen yang dibutuhkan yakni vektor gradien yang sesuai dengan persamaan (2.21) dan matriks heissian sesuai fungsi *likelihood* terhadap estimator β dan parameter dispersi α yang telah dijabar pada persamaan (4.9) dan (4.10). sedangkan matriks Hessian adalah turunan kedua dari fungsi local likelihood terhadap parameter β dan parameter α. Turunan kedua terhadap β dijabarkan sebagai berikut:

$$\frac{\partial^2 \ell(\beta, \alpha,)}{\partial \beta^2} = \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij} - x_{0j}) \left(- \frac{(\mathrm{x}_i).\exp(\mathrm{x}_i\beta).\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)}{\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)^2} \right.$$

$$\left. + \frac{\exp(\mathrm{x}_i\beta)(\mathrm{x}_i).\left(\alpha \exp\,(\mathrm{x}_i\beta)(\mathrm{x}_i)\right)}{\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)^2} + \frac{-y_i(\mathrm{x}_i).\alpha \exp(\mathrm{x}_i\beta)(\mathrm{x}_i)}{\left(1 + \alpha \exp(\mathrm{x}_i\beta)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} \left(- \frac{(\mathrm{x}_i).\exp(\mathrm{x}_i\beta).(\mathrm{x}_i) + (\mathrm{x}_i).(\mathrm{x}_i).\alpha\left(\exp(\mathrm{x}_i\beta)\right)^2}{\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)^2} \right.$$

$$\left. + \frac{-(\mathrm{x}_i).(\mathrm{x}_i).\alpha\left(\exp(\mathrm{x}_i\beta)\right)^2}{\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)^2} + \frac{-y_i(\mathrm{x}_i).\alpha\exp(\mathrm{x}_i\beta).(\mathrm{x}_i)}{\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)^2} \right)$$

$$= \sum_{i=j}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij}-x_{0j}) \left(- \frac{(\mathrm{x}_i).\exp(\mathrm{x}_i\beta).(\mathrm{x}_i) - y_i(\mathrm{x}_i).\alpha\exp(\mathrm{x}_i\beta).(\mathrm{x}_i)}{\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij}-x_{0j}) \left(\frac{-(\mathrm{x}_i).\exp(\mathrm{x}_i\beta).(\mathrm{x}_i)(y_i\alpha-1)}{\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)^2} \right) \tag{4.11}$$

Sedangkan turunan kedua terhadap α dijabarkan sebagai berikut:

$$\frac{\partial^2 \ell(\beta,\alpha)}{\partial\alpha^2}$$

$$= \sum_{i=1}^{n} \prod_{j=1}^{p} K_{hj}(x_{ij}-x_{0j}) \left\{ \psi'\left(y_i + \frac{1}{\alpha}\right) - \psi'\left(\frac{1}{\alpha}\right) \right.$$

$$+ \frac{-\left(y_i - \exp(\mathrm{x}_i\beta)\right)\left(1+\alpha\exp(\mathrm{x}_i\beta)\right) - \alpha y_i \exp(\mathrm{x}_i\beta) + \alpha\left(\exp(\mathrm{x}_i\beta)\right)^2}{\alpha^2\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)^2}$$

$$\left. + \frac{\left(\exp(\mathrm{x}_i\beta)\right)\alpha}{1+\alpha\exp(\mathrm{x}_i\beta)\alpha^3} - \frac{2\ln\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)}{\alpha^3} \right\} \tag{4.12}$$

Agar memaksimumkan fungsi *likelihood*, matriks heissian H harus definit negatif. Suatu matriks merupakan matriks definit negatif jika dan hanya jika bentuk kuadrat x'Ax bernilai negatif untuk setiap vektor x ≠ 0. Pengestimasian parameter β dan α dilakukan menggunakan metode *Locally Weigthed Maximum Likelihood Estimator* dengan iterasi *Newton Raphson* untuk mendapatkan penyelesaian fungsi likelihood yang tidak linier dalam parameter.

$$\beta_{(m+1)} = \beta_m - \left[\frac{\partial^2 \ell(\beta,\alpha,x_0)}{\partial\beta\partial\beta'}\right]_m^{-1} \left[\frac{\partial^2 \ell(\beta,\alpha,x_0)}{\partial\beta}\right]_m \qquad m = 0,1,2,\ldots \tag{4.13}$$

Dengan

$$\beta_{(m+1)} = \begin{bmatrix} \beta_{(m+1)} \\ \vdots \\ \beta_{(p+1)} \end{bmatrix}$$

$$\beta_m = \begin{bmatrix} \beta_{1m} \\ \vdots \\ \beta_{pm} \end{bmatrix}$$

$$g = \left[\frac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta}\right] = \begin{bmatrix} \dfrac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0} \\ \vdots \\ \dfrac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_p} \end{bmatrix}$$

$$H = \left[\frac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta \partial \beta'}\right] = \begin{bmatrix} \dfrac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0^2} & \cdots & \dfrac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0 \partial \beta_p} & \dfrac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0 \partial \alpha} \\ & \ddots & \vdots & \dfrac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_p \partial \alpha} \\ & & \dfrac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \beta_0^2} & \dfrac{\partial^2 \ell(\beta, \alpha, x_0)}{\partial \alpha^2} \end{bmatrix} \quad (4.14)$$

4.2 Pemodelan Regresi Binomial Negatif pada Kasus HIV/AIDS di Jawa Timur berdasarkan Estimator Lokal Linier

Jumlah kasus HIV/AIDS di Jawa Timur dilaporkan 17.394 orang dan 36.881 kasus HIV. Dari jumlah tersebut 3.679 (21,1%) diantaranya meninggal dunia. Angka tersebut sesungguhnya jauh lebih kecil dibandingkan angka yang sebenarnya terjadi dan dari hasil estimasi sampai dengan tahun 2012 diperkirakan jumlah ODHA di Jawa Timur mencapai 57.321 orang. Sejak September 2003, Provinsi di Jawa Timur ditetapkan sebagai wilayah dengan prevalensi HIV/AIDS yang terkonsentrasi bersama 5 provinsi lainnya, yaitu DKI Jakarta, Papua, Bali, Riau dan Jawa Barat.

Data yang diestimasikan pada model regresi binomial negatif menggunakan estimator lokal linier menggunakan data diskrit dengan variabel terikatnya adalah banyaknya penderita HIV/AIDS di Jawa Timur tahun 2018 dan variabel bebasnya adalah banyaknya pendonor darah dan banyaknya pengguna NAPZA. Berdasarkan data pada Lampiran 1 terdapat 38 Kabupaten/Kota di Jawa Timur yang akan diestimasi. Sebelum mengestimasi data tersebut, terlebih dahulu dilakukan analisis statistika deskriptif yang bertujuan untuk mengidentifikasi data tersebut.

Tabel 4.1 Staistik Deskriptif

Variabel	Mean	Variance	Minumum	Maximum
Y	26,38	1555,98	2	161
X1	13383	1354756	1011	48222
X2	130,3	50722,7	3	878

Berdasarkan Tabel 4.1 dapat dilihat nilai rata-rata penderita penyakit HIV/AIDS di Jawa Timur tahun 2018 sebesar 26,38 dengan paling sedikitnya penderita berjumlah 2 dan penderita paling banyak 161 orang pada salah satu daerah. Variabel x_1 adalah banyaknya pendonor darah yang memiliki rata-rata sebesar 13383 dengan jumlah paling sedikit 1011 dan paling banyak 48222 pendonor dari salah satu daerah. Variabel x_2 adalah banyaknya pengguna NAPZA dengan rata-rata 130,3 dengan nilai minum atau paling rendahnya 3 dan tertingginya 878 orang pada salah satu daerah. Nilai variansi pada variabel respon y (banyak penderita HIV/AIDS bernilai jauh lebih besar dari nilai rata-rata, hal tersebut mengidentifikasikan bahwa terjadi *overdispersion* pada data. Indikasi tersebut dikonfirmasi dengan pengujian statistik *Likelihood Ratio Test* (LRT) dengan hipotesis yang digunakan sebagai berikut:

H_0 : $LR = \chi^2$ (banyaknya penderita HIV/AIDS mengalami *equidispersion*)

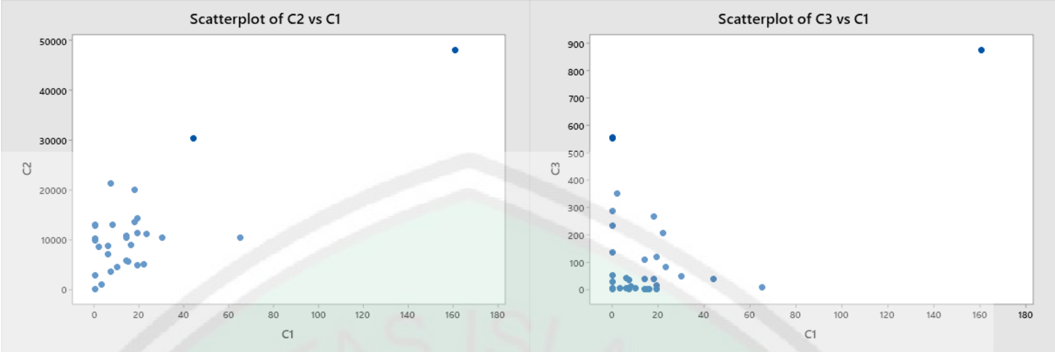
H_1 : $LR > \chi^2$ (banyaknya penderita HIV/AIDS mengalami *overdispersion*)

Tabel 4.2 Hasil Likelihood Ratio Test

Number of obs	16
LR chi2(2)	32,85
Prob > chi2	0,0000
Pseudo R2	0,0764

Berdasarkan output STATA pada Tabel 4.2 dapat diketahui bahwa data tersebut mengalami dispersi pada nilai rata-rata dengan nilai *p-value* LRT tersebut sebesar 0,0000 serta nilai Likelihood Ratio sebesar 32,85 sedangkan nilai χ^2 sebesar 24,99. Sehingga dapat diketahui bahwa nilai *Likelihood Ratio* > χ^2 yaitu 32,85 > 24,99, dimana nilai α yaitu 0,05 maka *p-value* < α yaitu 0,00 < 0,05. Sehingga dapat ditarik kesimpulan bahwa data tersebut terjadi overdispersion.

Gambar berikut adalah gambar plot antara variable respon dan 2 variabel prediktor:



Gambar 4.1 *Scatterplot* antara Variabel Respon dengan Variabel Prediktor

Berdasarkan Gambar 4.1, dapat diasumsikan bahwa variabel x_i dan x_2 memiliki pola yang tidak beraturan terhadap variable respon y atau dapat dikatakan tidak mengikuti suatu pola tertentu sehingga pendekatan yang cocok untuk pemodelan data tersebut dengan menggunakan pendekan regresi nonparametrik.

Estimasi model regresi nonparametrik berdasarkan estimator lokal linier membutuhkan nilai *bandwidth* optimal yang akan digunakan pada pembentukan matriks K_h. Pada penelitian ini, penentuan *bandwidth (h)* optimal dilakukan berdasarkan kriteria maksimum MLCV(h).

Tabel 4.3 Penentuan *bandwidth optimal* pada variable *predictor* 1

Bandwidth	MLCV(h)
200	29.5407
200	29.5406
199	29.5407
200	**29.5410**
199	29.5408
198	29.5408
199	29.5409
198	29.5409
200	29.5408

197	29.5408

Tabel 4.4 Penentuan *Bandwidth Optimal* pada Variabel Prediktor 2

Bandwidth	MLCV(h)
10	37.3702
9	37.4469
7	40.4085
9	45.1828
10	46.2983
10	48.5534
10	49.3653
10	49.7446
8	**52.8617**
10	52.6006

Berdasarkan tabel di atas didapatkan *bandwidth* optimal variabel prediktor 1 sebesar 200 dengan nilai MLCV(h) maksimum yaitu 29,5410 dan pada prediktor 2 yaitu sebesar 8 dengan MLCV(h) maksimum 52,8617 .

Setelah memperoleh nilai *bandwidth* optimal pada masing-masing variabel prediktor, selanjutnya dilakukan estimasi model regresi Binomial Negatif dengan pendekatan regresi nonparametrik berdasarkan estimator lokal linier. Sehingga diperoleh nilai parameter β.

Tabel 4.5 Nilai Parameter

Parameter	Nilai Parameter
β_0	6,46322
β_1	0,000273
β_2	0,000714
β_3	0,002765

Berdasarkan tabel 4.5 tersebut dapat diperoleh model umum seperti berikut:

$$\hat{y} = \exp(\, 6{,}46322 + 0{,}000273x_1 + 0{,}000714(x_1 - 29.5410) + 0{,}002765\ (x_2 - 52.8617)$$

$$= \exp(6{,}29602 + 0{,}000987x_1 + 0{,}002765x_2) \quad (4.14)$$

$\hat{y}_3 =$ exp(6,29602+(0,000987(8491)+0,002765(352))

$= 1{,}9876$

Berdasarkan model tersebut, diperoleh hasil estimasi banyak penderita HIV di Kabupaten Trenggalek sebesar 1,98 atau 2. Nilai yang diperoleh sama dengan data riil banyak penderita HIV Kabupaten Trenggalek. Interpretasi yang diperoleh dengan menggunakan persamaan 4.14 apabila terjadi kenaikan variabel banyak pendonor (x_1) sebesar 8491 dan variabel banyak pengguna NAPZA (x_2) sebesar 357, maka variabel respon akan mengalami perubahan 0,1 kali dari semula.

Uji kesesuaian model atau biasa disebut dengan *goodness of fit* digunakan untuk menguji apakah model tersebut sudah sesuai atau belum. Uji yang digunakan adalah uji statistik *Deviance*. Nilai *Deviance* yang diperoleh adalah sebagai berikut:

Tabel 4.6 Kriteria Kesesuaian Model

Uji Kesesuaian Model	Nonparametrik
Nilai *Deviance*	8,2244
Nilai *P-value*	0,8897

Hipotesis yang digunakan untuk menguji kesesuaian model menggunakan nilai *deviance* adalah:

H_0 : Model telah sesuai

H_1 : Model tidak sesuai

Berdasarkan Tabel 4.6 dapat diketahui nilai statistik *uji deviance* regresi binomial negative dengan pendekatan regresi nonparametrik sebesar 8,2244 dan nilai p-value sebesar 0,8897 dengan α sebesar 0,05 sedangkan nilai χ^2 sebesar 24,99. Sehingaga keputusan terima H_0 dimana *p-value* $> \alpha$ atau D < x^2 maka model tersebut telah sesuai.

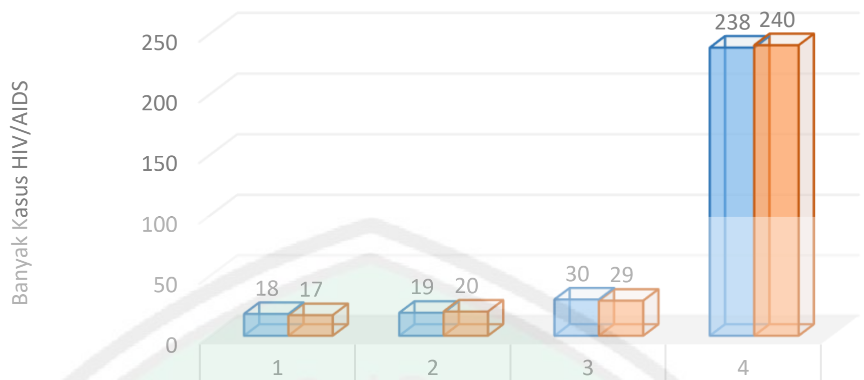
Langkah selanjutnya diterapkan atau dicoba pada data testing pada model binomial megatif menggunakan estimator lokal data outsampel. Data yang digunakan sebagai data outsampel yaitu 20% data dari 20 data yang diambil secara acak sebanyak 4 data.

Tabel 4.7 Data *outsampel*

No	Kota/Kab	Y	$\hat{y}$	*Absolute Percentage error*
1	Tulungagung	18	17	0,894%
2	Probolinggo	19	20	0,314%
3	Mojokerto	30	29	0,306%
4	Surabaya	238	240	0,354%

Berdasarkan Tabel 4.7 diketahui bahwa y observasi pada Kabupaten Tulungagung sebanyak 18 kasus dengan banyaknya x_1 dan x_2 adalah 20098 kasus dan 37 kasus menghasilkan y estimasi sebesar 17 kasus dengan error sebesar 0,894. Kabupaten Probolinggo dengan banyak 19 kasus dan menghasilkan y estimasi sebesar 20 dengan error sebesar 0,314. Kabupaten Mojokerto dengan banyak 30 kasus dan menghasilkan y estimasi sebesar 29 dengan error sebesar 0,306. Kota Surabaya memiliki 238 kasus dengan y estimasi sebesar 240 dan error sebesar 0,354. Berdasarkan error dari masing-masing kabupaten atau kota diperoleh nilai MAPE sebesar 4,387%. Menurut Kristien Margie (2015) apabila nilai MAPE yang dihasilkan kurang dari 10%, maka dapat dikatakan bahwa persentase penyimpangan antara data actual atau observasi dengan data estimasi memiliki pendugaan yang sangat baik.

Diagram antara observasi dengan hasil estimasi pada variabel respon banyak penderita HIV/AIDS di Jawa Timur :



	1	2	3	4
Data Observasi	18	19	30	238
Data Estimasi	17	20	29	240

Gambar 4.2 Plot Observasi dan Estimasi

Berdasarkan Gambar 4.2 tersebut dapat diketahui bahwa hasil data observasi dengan data estimasi memiliki selisih yang tidak jauh, hal ini menunjukkan bahwa adanya kesesuaian antara hasil observasi dengan hasil estimasi pada kasus HIV/AIDS di Jawa Timur.

BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan maka diperoleh kesimpulan sebagai berikut:

1. Estimasi model regresi binomial negatif dengan menggunakan pendekatan regresi nonparametrik multiprediktor adalah sebagai berikut:

$$\ell(\beta,\alpha)=\sum_{i=1}^{n}\prod_{j=1}^{p}K_{hj}\left(x_{ij}-x_{0j}\right)\Bigg\{\ln\Gamma\left(y_i+\frac{1}{\alpha}\right)-\ln y_i!-\ln\Gamma\left(\frac{1}{\alpha}\right)$$
$$-\frac{1}{\alpha}\ln\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)+y_i\ln\alpha+y_i(\mathrm{x}_i\beta)$$
$$-y_i\ln\left(1+\alpha\exp(\mathrm{x}_i\beta)\right)\Bigg\}$$

   Pengestimasian parameter β dan α dilakukan menggunakan metode *Locally Weigthed Maximum Likelihood Estimator* dengan iterasi *Newton Raphson* untuk mendapatkan penyelesaian fungsi *likelihood* yang tidak linier dalam parameter.

$$\beta_{(m+1)}=\beta_m-\left[\frac{\partial^2\ell(\beta,\alpha,x_0)}{\partial\beta\partial\beta'}\right]_m^{-1}\left[\frac{\partial^2\ell(\beta,\alpha,x_0)}{\partial\beta}\right]_m \qquad m=0,1,2,\ldots$$

2. Model regresi binomial negatif menggunakan estimator lokal linier pada kasus HIV/AIDS di Jawa Timur menggunakan bandwidth (h) optimal yaitu 29,5410 dan 52,8617, menghasilkan model sebagai berikut:

$$\hat{y}=\exp(6{,}29602+0{,}000987x_1+0{,}002765x_2)$$

5.2 Saran

Berdasarkan hasil penelitian, saran yang dapat diberikan oleh peneliti adalah sebagai berikut:

1. Mengembangkan estimasi model regresi binomial negatif nonparametrik dengan kasus birespon. Penelitian ini masih sangat terbatas pada estimatornya, sehingga dapat dikembangkan dengan estimator-estimator yang lainnya.
2. Pemerintah Provinsi Jawa Timur diharapkan lebih memberikan perhatian khusus terhadap kabupaten/kota yang memiliki banyak kasus HIV/AIDS dan

diharapkan mampu mengendalikan faktor-faktor yang dapat menambah banyak kasus HIV/AIDS.

DAFTAR PUSTAKA

Astuti, E.T., dan Yanagawa, T,. 2002, Testing Trend for Count Data with Extra Poisson Variability, Biometrics, 58, 398-402.

Barry, S. C. dan Welsh, A.H,. 2002. Generalized Additive modeling and Zero Inflated Count Data. Ecological Modelling, 157. 179-188

Cameron, A.C., dan Trivedi, P.K., 1998. Regression Analysis of Count Data . New York: Cambridge University Press

Dinas Kesehatan Provinsi Jawa Timur, 2017, Profil Kesehatan Provinsi Jawa Timur, Dinas Kesehatan Provinsi Jawa Timur, Surabaya.

Eubank, R.L., 1999, *Nonparametric Regression and Spline* Smoothing 2nd Edition, New York: Marcel Deker.

Fan, J.,Gijbels, I, Chung-Hu,T., dan Huang, L, 1996, A Study of Variable Bandwidth Selection For Local Polynomial Regression, Journal of Statistica Sinica, 6(1996), 113-127.

Hardle, W., Miller, M., Sperlich, S. Dan Werwatz,A., 2004, *Nonparametric and semiparametric Models*, New York: Spinger.

Hilbe, M.J.,2011, Negative Binomial Regression 2nd Edition. New York: Cambridge University Press

Liang, H dan Cheng, D.Z., 2005. Assement of Esophageal Pressure in Gastreeshopageal Relux Disease by *Local Regression Annalysis of Biomedical engineering*. 33: 847-853

Nottingham, Q.J dan Cook D.F, 2001, *Local Linier Regression for estimating time series data, Journal of Computational Statistics and Data Analysis*, 37, 209-217.

Park, B.J dan Lord, D.,2008, *Adjusment for The Maximum Likelihood Estimate of The Negative Binomial Dispersion Parameter*, USA: Texas University

LAMPIRAN

Lampiran 1

No	Kabupaten	HIV	Pendonor darah	Pengguna Napza
1.	Pacitan	15	5530	0
2.	Ponorogo	0	0	27
3.	Trenggalek	2	8491	352
4.	Tulungagung	18	20098	37
5.	Blitar	22	5097	206
6.	Kediri	0	13019	233
7.	Malang	0	0	135
8.	Lumajang	14	10850	110
9.	Jember	0	0	555
10.	Banyuwangi	0	0	288
11.	Bondowoso	6	7122	40
12.	Situbondo	14	5779	37
13.	Probolinggo	19	4839	118
14.	Pasuruan	65	10505	7
15.	Sidoarjo	44	30464	38
16.	Mojokerto	30	10338	47
17.	Jombang	7	21371	36
18.	Nganjuk	18	13510	268
19.	Madiun	6	8723	3
20.	Magetan	0	2794	3
21.	Ngawi	16	8912	0
22.	Bojonegoro	8	12987	10
23.	Tuban	23	11148	81
24.	Lamongan	0	12828	0
25.	Gresik	0	0	53
26.	Bangkalan	0	9878	0
27.	Sampang	10	4564	3
28.	Pamekasan	0	0	3
29.	Sumenep	3	1011	6

Kota/*Municipality*				
30	Kediri	0	0	558
31	Blitar	0	10149	6
32	Malang	161	48222	878
33	Probolinggo	14	10436	0
34	Pasuruan	7	3527	1
35	Mojokerto	19	11358	0
36	Madiun	19	14274	15
37	Surabaya	238	43691	31
38	Batu	0	0	6

Lampiran 2

Fitting Poisson model:

Iteration 0: log likelihood = -1128.158
Iteration 1: log likelihood = -476.66948
Iteration 2: log likelihood = -278.93893
Iteration 3: log likelihood = -274.57206
Iteration 4: log likelihood = -274.56596
Iteration 5: log likelihood = -274.56596
Fitting constant-only model:
Iteration 0: log likelihood = -154.58259
Iteration 1: log likelihood = -136.36746
Iteration 2: log likelihood = -136.36334
Iteration 3: log likelihood = -136.36334
Fitting full model:
Iteration 0: log likelihood = -131.08421
Iteration 1: log likelihood = -126.52444
Iteration 2: log likelihood = -125.95072
Iteration 3: log likelihood = -125.93872
Iteration 4: log likelihood = -125.93871

Negative binomial regression Number of obs = 38
LR chi2(2) = 32.85
Dispersion = mean Prob > chi2 = 0.0000
Log likelihood = -125.93871 Pseudo R2 = 0.0764

A	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
B	.0001144	.0000299	3.83	0.000	.0000558	.0001729
C	-.002419	.0015519	-1.56	0.119	-.0054606	.0006226
_cons	1.418969	.3369253	4.21	0.000	.758608	2.079331
/lnalpha	.6026594	.2880277			.0381355	1.167183
alpha	1.826971	.5262182			1.038872	3.21293

Likelihood-ratio test of alpha=0: chibar2(01) = 29.25 Prob>=chibar2 = 0.000

Lampiran 3

TABEL BANDWIDTH OPTIMAL PRED 1

h1	MLCV
200	29.5407
300	29.5406
199	29.5407
200	29.5407
199	29.5408
198	29.5408

TABEL BANDWIDTH OPTIMAL PRED 2

h2	MLCV
10	37.3702
9	37.4469
7	40.4085
9	45.1828
10	46.2983
10	48.5534

199 29.5409 10 49.3653

198 29.5409 10 49.7446

200 29.5408 8 52.8617

197 29.5408 10 52.6006

bandwidth optimal pred 1 = 200 bandwidth optimal prediktor 2 = 8

ESTIMASI BETA PARAMETRIK=

Call: glm.nb(formula = y ~ x11 + x12, data = datanew, init.theta = 3.026401879, link = log)

Coefficients:

(Intercept) x11 x12

4.03746764 0.00004666 0.00807260

Degrees of Freedom: 13 Total (i.e. Null); 11

Residual Null Deviance:20.03 Residual Deviance: 14.84

AIC: 160.6

----

yasli	y nonpar	error nonpar
2	1	0.756
22	18	0.9527
16	17	1.3689
14	16	0.344
19	20	0.314
65	62	0.834
30	29	0.3064
17	20	0.593
18	17	0.894
16	17	0.192
18	16	0.348
23	25	0.5145
10	11	0.404
13	14	0.977
19	20	0.234
238	240	0.3540

DEV_NONPAR PVALNONPAR

8.2244496 0.8897655

Nilai R2 Nonparametrik : 0.8782222

parameter nonparametrik=

6.46322   0.0002732   0.027143   0.276522

Lampiran 4

TABEL NILAI KRITIS DISTRIBUSI CHI-SQUARE

df	0,1	0,05	0,025	0,001	0,005
1	2,705543	3,841459	5,023886	6,634897	7,879439
2	4,605170	5,991465	7,377759	9,210340	10,596635
3	6,251389	7,814728	9,348404	11,344867	12,838156
4	7,779440	9,487729	11,143287	13,276704	14,860259
5	9,236357	11,070498	12,832502	15,086272	16,749602
6	10,644641	12,591587	14,449375	16,811894	18,547584
7	12,017037	14,067140	16,012764	18,475307	20,277740
8	13,361566	15,507313	17,534546	20,090235	21,954955
9	14,683657	16,918978	19,022768	21,665994	23,589351
10	15,987179	18,307038	20,483177	23,209251	25,188180
11	17,275009	19,675138	21,920049	24,724970	26,756849
12	18,549348	21,026070	23,336664	26,216967	28,299519
13	19,811929	22,362032	24,735605	27,688250	29,819471
14	21,064144	23,684791	26,118948	29,141238	31,319350
15	22,307130	24,995790	27,488393	30,577914	32,801321
16	23,541829	26,296228	28,845351	31,999927	34,267187
17	24,769035	27,587112	30,191009	33,408664	35,718466
18	25,989423	28,869299	31,526378	34,805306	37,156451
19	27,203571	30,143527	32,852327	36,190869	38,582257
20	28,411981	31,410433	34,169607	37,566235	39,996846
21	29,615089	32,670573	35,478876	38,932173	41,401065
22	30,813282	33,924438	36,780712	40,289360	42,795655
23	32,006900	35,172462	38,075627	41,638398	44,181275
24	33,196244	36,415029	39,364077	42,979820	45,558512
25	34,381587	37,652484	40,646469	44,314105	46,927890
26	35,563171	38,885139	41,923170	45,641683	48,289882
27	36,741217	40,113272	43,194511	46,962942	49,644915
28	37,915923	41,337138	44,460792	48,278236	50,993376
29	39,087470	42,556968	45,722286	49,587884	52,335618
30	40,256024	43,772972	46,979242	50,892181	53,671962
31	41,421736	44,985343	48,231890	52,191395	55,002704
32	42,584745	46,194260	49,480438	53,485772	56,328115
33	43,745180	47,399884	50,725080	54,775540	57,648445
34	44,903158	48,602367	51,965995	56,060909	58,963926
35	46,058788	49,801850	53,203349	57,342073	60,274771
36	47,212174	50,998460	54,437294	58,619215	61,581179
37	48,363408	52,192320	55,667973	59,892500	62,883335
38	49,512580	53,383541	56,895521	61,162087	64,181412
39	50,659770	54,572228	58,120060	62,428121	65,475571
40	51,805057	55,758479	59,341707	63,690740	66,765962

RIWAYAT HIDUP

Nuchaila Ainiyah lahir di Malang pada 1 Agustus 1998. Memiliki nama panggilan Nucha. Alamatnya berada di Dsn.Gedangsewu RT/RW 03/08 Ds.Randuagung, Kec.Singosari Kab.Malang. Merupakan anak pertama dari Bapak Choidin dan Ibu Khoirul Abidah.

Pendidikan yang pernah ditempuh yaitu TK Islam Kartini. Kemudian melanjutkan sekolahnya di MI Al-Ma'arif 09 dan lulus pada tahun 2010. Menempuh pendidikan SMP di MTs. Darul Karomah Randuagung dan lulus pada tahun 2013. Melanjutkan pendidikan SMA di SMA Al-Rifa'ie Gondanglegi lulus pada tahun 2016.

Tahun 2016 melanjutkan studi ke jenjang pendidikan strata 1 di Universitas Islam Negeri Maulana Malik Ibrahim Malang mengambil jurusan Matematika Fakultas Sains dan Teknologi. Aktif mengikuti kegiatan organisasi yang ada di dalam kampus, seperti menjadi Pengurus HMJ Matematika UIN Malang (2016-2017) dan (2017-2018), Pengurus DEMA Fakultas Sains dan Teknologi UIN Malang (2018-2019)..

KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAUALANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No. 50 Dinoyo Malang Telp./Fax.(0341)558933

BUKTI KONSULTASI SKRIPSI

Nama : Nuchaila Ainiyah
NIM : 16610008
Fakultas/Jurusan : Sains dan Teknologi / Matematika
Judul Skripsi : Regresi Binomial Negatif pada Kasus HIV/AIDS di Jawa Timur Berdasarkan Estimator Lokal Linier
Pembimbing I : Ria Dhea Layla N.K, M.Si
Pembimbing II : Dr. Hairur Rahman, M.Si

No	Tanggal	Hal	Tanda Tangan
1	7 Oktober 2020	Setor dan Konsultasi Judul	1.
2	6 Februari 2020	Konsultasi Bab I	2.
3	14 Februari 2020	Revisi Bab I dan Konsultasi Bab II	3.
4	23 Februari 2020	Revisi Bab II dan Konsultasi Bab III	4.
5	26 Februari 2020	Konsultasi dan Revisi Kajian Agama	5.
6	2 Maret 2020	Konsultasi dan Revisi Kajian Agama	6.
7	4 Maret 2020	ACC Kajian Agama	7.
8	14 Maret 2020	ACC Bab I, Bab II dan Bab III	8.
9	19 Oktober 2020	Konsultasi Bab IV	9.
10	25 Oktober 2020	Revisi Bab IV dan Konsultasi Bab V	10.
11	18 November 2020	ACC Bab IV dan Bab V	11.
12	2 Desember 2020	Revisi semua bab pasca siding	12.

13	31 Desember	ACC keseluruhan	13.

Malang, 31 Desember 2020
Mengetahui,
Ketua Jurusan Matematika

Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001