

# Analisis *Propensity Score* menggunakan Regresi Logistik pada Kasus Data HIV/AIDS LSM ORBIT Surabaya

<sup>1</sup>Farida Islamiah dan <sup>2</sup>Bambang Widjanarko Otok

<sup>1,2</sup>Statistika, FMIPA, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: <sup>1</sup>farida\_paser@yahoo.com dan <sup>2</sup>[bambang\\_wo@statistika.its.ac.id](mailto:bambang_wo@statistika.its.ac.id)

**Abstrak**—*Propensity score* merupakan probabilitas bersyarat mendapatkan perlakuan tertentu dengan melibatkan kovariat yang diamati. Metode ini digunakan untuk mengurangi bias dalam estimasi dampak dari perlakuan pada data yang bersifat observasi karena adanya faktor *confounding*. Dalam penelitian observasional, kovariat biasanya tidak seimbang antara kelompok perlakuan dan kelompok kontrol. Metode yang digunakan dalam *propensity score* yaitu stratifikasi, *matching* atau pembobotan untuk menghilangkan *confounding*. Jika perlakuan bersifat biner, model regresi logistik dan probit dengan variabel basis kovariat dan nilai prediksi dari regresi sebagai *propensity score*. Pada penelitian ini, data yang digunakan yaitu data sekunder hasil survei yang dilakukan oleh LSM ORBIT Surabaya mengenai faktor-faktor yang mempengaruhi HIV/AIDS pada pengguna napza suntik (Penasun). Variabel selalu memakai kondom sebagai variabel *confounding* dimana perlakuannya adalah selalu menggunakan kondom, sedangkan variabel respon yaitu status HIV/AIDS. Karena variabel *confounding* bersifat biner, maka *propensity score* menggunakan regresi logistik. Metode yang digunakan dalam *propensity score* yaitu pembobotan pada unit observasi. Estimasi *propensity score* menggunakan regresi logistik dengan metode *Maximum Likelihood Estimation* (MLE), karena hasil estimasi tidak *close form*, maka dilakukan iterasi Newton-Raphson. Hasil analisis *propensity score* dengan pembobot pada kasus HIV/AIDS yaitu pengguna napza suntik yang selalu menggunakan kondom dapat terkena HIV/AIDS 1,25 kali dibandingkan dengan tidak selalu menggunakan kondom.

**Kata Kunci**—HIV/AIDS, pembobotan, *propensity score*, dan regresi logistik.

## I. PENDAHULUAN

Metode *propensity score* pertama kali diperkenalkan oleh Rosenbaum dan Rubin pada tahun 1983. Metode ini digunakan untuk mengurangi bias dalam estimasi dampak dari perlakuan pada data yang bersifat observasi. Dalam studi observasi unit perlakuan dan kontrol tidak acak, estimasi atas perlakuan akan bias oleh adanya faktor *confounding*. Bias tersebut akan tereduksi ketika hasil perbandingan antara kelompok perlakuan dan kontrol hampir sama [1].

Dalam penelitian observasional, kovariat biasanya tidak seimbang antara kelompok perlakuan dan kelompok kontrol. Rujukan [2] telah menunjukkan bahwa mengamati kovariat yang seimbang pada setiap nilai *propensity score*; hal tersebut menunjukkan kelompok perlakuan dan kontrol dengan *propensity score* yang sama memiliki distribusi yang sama dari kovariat yang diamati [3].

Metode berbasis *propensity score* bertujuan untuk menghilangkan ketidakseimbangan ini. Ada empat metode berbasis *propensity score* yaitu stratifikasi, *matching*, *covariate adjustment* dan *inverse probability weighting by propensity score*. Stratifikasi dan pencocokan adalah metode yang umum digunakan dan bertujuan untuk menjadikan data dalam percobaan terkontrol acak dimana perbandingan dari kelompok perlakuan bermakna. *Covariate adjustment* oleh *propensity score* juga pendekatan yang sering digunakan karena mudah dalam penerapannya. Pendekatan yang jarang digunakan yaitu *inverse probability weighting by propensity score* [4]. *Propensity score* pada metode bobot sering digunakan pada regresi terbobot [5]. Metode yang populer yang digunakan untuk estimasi perbedaan rata-rata perlakuan ditunjukkan oleh [6], dimana unit observasi distratifikasikan berdasarkan estimasi *propensity score* dan perbedaan estimasi sebagai rata-rata antar pengaruh strata. Pendekatan alternatif yaitu menambahkan *confounding* dalam estimasi *propensity score* untuk pembobotan pada unit observasi [7].

Beberapa penelitian tentang metode yang digunakan untuk estimasi *propensity score* yaitu, D'Agostino [8] menggunakan analisis diskriminan untuk estimasi *propensity score*, McCaffrey [9] menggunakan model *generalized boosted*, Setouguchi [10] menggunakan studi simulasi pada *neural network* dan *classification tree*. Sebagian besar metode tersebut memiliki kelemahan seperti kompleksitas dan permasalahan dalam interpretasi, serta menghadapi banyak kerumitan dari hasil algoritma yang kompleks dan implementasinya. Jika perlakuan bersifat biner, model regresi logistik dan probit dengan variabel basis kovariat dan nilai prediksi dari regresi sebagai *propensity score* [11]. Oleh karena itu, regresi logistik banyak digunakan untuk estimasi *propensity score* karena mudah dalam interpretasinya [3].

Regresi logistik adalah teknik statistik yang estimasi probabilitas dari suatu peristiwa (variabel respon) berdasarkan faktor-faktor yang diketahui (variabel prediktor) yang diekspektasi mempengaruhi terjadinya peristiwa (perlakuan dalam kasus perhitungan *propensity score*). Variabel respon diasumsikan dua nilai yaitu terjadi atau tidak. Hasil regresi logistik adalah probabilitas (mulai dari 0 hingga 1). Hasil regresi logistik dapat digunakan untuk menghitung nilai *propensity score* [3]. Model regresi logistik yang sering digunakan dalam penelitian observasional untuk menilai hubungan antara perlakuan dan hasil dikotomis (variabel

respon), sedangkan *confounders* untuk membandingkan antara kelompok dan untuk mengurangi bias [12].

Penambahan kasus penderita HIV/AIDS yang disumbangkan oleh pengguna Narkotika, Psikotropika, Alkohol dan Penggunaan Zat Additive (Napza) suntik cukup besar. Departemen Kesehatan (Depkes) melaporkan bahwa sumbangan pengguna Napza suntik (Penasun) terhadap semua kasus HIV di Indonesia sebesar 19,9%–22,1% sampai tahun 2001 dan 2003. Salah satu penyebabnya adalah angka prevalensi HIV pada kalangan penasun yang meningkat tajam, sampai tahun 2002 angka tersebut bervariasi antara 10%–80% [1]. Penularan virus mematikan ini pada kalangan penasun mengalami penurunan yang cukup signifikan. Hasil Survei Terpadu Biologis dan Perilaku (STBP) 2007 dan 2011 menemukan bahwa prevalensi HIV pada kelompok penasun mencapai 52,4 persen. Yang mengkhawatirkan, penasun ini sering melakukan hubungan seks berisiko yang dapat menularkan HIV ke populasi umum. Penelitian terkait Napza telah dilakukan oleh Besral [1], Praptoraharjo [13], Pratiwi [14], Arniti [15] dan Kristanti [16].

Berdasarkan uraian di atas, maka penelitian ini bertujuan untuk mengkaji *propensity score* dengan menggunakan metode regresi logistik pada kasus HIV/AIDS pada salah satu kota yang memiliki jumlah terinfeksi HIV/AIDS terbanyak yaitu Surabaya.

## II. TINJAUAN PUSTAKA

### A. Propensity Score

*Propensity score* untuk  $i = 1, 2, \dots, n$  sebagai probabilitas bersyarat bergantung pada unit perlakuan ( $Z_i = 1$ ) dibandingkan unit kontrol ( $Z_i = 0$ ) dengan vektor kovariat yang diamati  $x_i$ :

$$e(x_i) = P(Z_i = 1 | X_i = x_i) \quad (1)$$

Langkah pertama yang dilakukan dalam *propensity score* yaitu memilih kovariat sebagai *confounder* untuk estimasi *propensity score*. Proses pemilihan *confounder* dapat berdasarkan teori dan menunjukkan hubungan antara variabel [17]. Uji *chi square*, digunakan untuk memeriksa hubungan antar variabel, dengan hipotesis sebagai berikut:

$H_0$ : Tidak terdapat hubungan yang signifikan antar variabel

$H_1$ : Terdapat hubungan yang signifikan antar variabel

Taraf signifikansi:  $\alpha = 5\%$

Statistik uji:

$$\chi^2 = \sum_{r=1}^R \sum_{c=1}^C \frac{(n_{rc} - e_{rc})^2}{e_{rc}}; e_{rc} = \frac{n_r \cdot n_c}{n} \quad (2)$$

$r = 1, 2, \dots, R$ ;  $c = 1, 2, \dots, C$

dengan

$n_{rc}$  : frekuensi pengamatan sel ke- $rc$

$e_{rc}$  : frekuensi harapan sel ke- $rc$

Daerah kritik:  $H_0$  ditolak jika  $\chi^2 > \chi^2_{1-\alpha}$ ;  $df = (r-1)(c-1)$  atau  $p\text{-value} < \alpha$  [18].

### B. Regresi Logistik

Regresi logistik digunakan jika variabel respon bersifat kategorik (nominal atau ordinal) dengan variabel-variabel prediktor kontinu maupun kategorik [19]. Variabel respon  $Y$  yang bersifat random dan dikotomus, yakni bernilai 1 dengan probabilitas  $\pi$  dan bernilai 0 dengan probabilitas  $1-\pi$  disebut sebagai *point-binomial* [20].

Model regresi logistik dengan  $k$  variabel prediktor adalah [20]:

$$\pi(x) = \frac{\exp\left(\beta_0 + \sum_{j=1}^k \beta_j x_j\right)}{1 + \exp\left(\beta_0 + \sum_{j=1}^k \beta_j x_j\right)} \quad (3)$$

Persamaan (3) juga dapat ditulis sebagai,

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)} \quad (4)$$

Jika model pada persamaan (4) ditransformasi dengan menggunakan transformasi logit, maka akan menghasilkan bentuk logit

$$g(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \quad (5)$$

yang merupakan fungsi linier dalam parameter-parameternya.

Pengujian signifikansi parameter model dengan satu variabel prediktor dilakukan untuk mengetahui ada atau tidaknya hubungan antara suatu variabel prediktor dan variabel respon [20]. Langkah pengujian hipotesisnya adalah sebagai berikut.

Hipotesis:

$H_0$ :  $\beta_j = 0$ ;  $j = 1, 2, \dots, k$

$H_1$ :  $\beta_j \neq 0$

Statistik Uji (Le, 1998):

$$\text{Wald}(W) = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (6)$$

Rasio yang dihasilkan dari persamaan (6), dibawah hipotesis  $H_0$ , akan mengikuti distribusi normal baku [21]. Sehingga untuk memperoleh keputusan, nilai statistik uji dibandingkan dengan distribusi normal baku ( $Z$ ). Kriteria penolakan  $H_0$  adalah jika

$$|W| > Z_{\alpha/2}$$

### C. Propensity Score menggunakan Regresi Logistik

*Propensity score* umumnya diestimasi dengan menggunakan metode regresi logistik, analisis diskriminan dan klasifikasi pohon. Regresi logistik merupakan metode yang sering digunakan untuk estimasi *propensity score* [17]. Berdasarkan persamaan (1), *propensity score* menggunakan model regresi logistik, dengan variabel respon adalah biner dimana  $Z_i = 1$  untuk perlakuan dan  $Z_i = 0$  untuk unit kontrol dengan model sebagai berikut [3]:

$$e(x_i) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)} \quad (7)$$

dengan:

$\beta_0$  : konstanta

$\beta_1, \beta_2, \dots, \beta_k$  : koefisien regresi

$x_1, x_2, \dots, x_k$  : variabel prediktor.

#### D. Propensity Score Terbobot

*Propensity score* terbobot dilakukan ketika estimasi pengaruh perlakuan [9]. Untuk estimasi  $E(y_0|z=1)$ , misalkan pengamatan  $i$  pada sampel pembanding memiliki bobot  $w_i = e(\mathbf{x}_i)/(1 - e(\mathbf{x}_i))$ , kemungkinan bahwa pengamatan yang dipilih secara random dengan  $\mathbf{x}$  termasuk dalam perlakuan. Pengamatan  $y_i = y_{1i}$  jika pengamatan  $i$  kelompok perlakuan dan  $y_i = y_{0i}$  jika pengamatan  $i$  kelompok kontrol.

#### E. Evaluasi Propensity Score

Metode *propensity score* tidak dapat memasukkan kovariat yang tidak dapat diukur yaitu tidak memiliki korelasi dengan kovariat observasi. Meskipun demikian, kualitas dari memasukkan kovariat observasi pada *propensity score* terbobot mudah untuk dievaluasi. Estimasi *propensity score* terbobot akan sama dengan distribusi dari kasus. Hal ini menyatakan bahwa statistik pembobot pada kovariat pada kelompok kontrol akan sama dengan kelompok perlakuan. Salah satu cara untuk menilai kualitas *propensity score* terbobot yaitu membandingkan variasi statistik seperti rata-rata, median, variansi dan statistik Kolmogorov Smirnov (KS) pada setiap kovariat [9].

Pengujian KS untuk kelompok perlakuan dan kelompok kontrol digunakan karena uji ini sensitif terhadap dua perbedaan yang ada antara dua distribusi. Uji KS untuk dua kelompok ini dikembangkan oleh Smirnov. Uji ini juga memakai nama Kolmogorov karena kemiripannya dengan uji satu kelompok yang dikembangkan oleh Kolmogorov. Hipotesis pada pengujian KS dua kelompok dengan uji dua sisi, yaitu

$$H_0 : F_T(\hat{e}(x)) = F_C(\hat{e}(x)) \text{ untuk semua } \hat{e}(x)$$

$$H_1 : F_T(\hat{e}(x)) \neq F_C(\hat{e}(x)) \text{ untuk paling tidak satu } \hat{e}(x)$$

Jika  $S_T(x)$  dan  $S_C(x)$  berturut-turut adalah fungsi-fungsi distribusi kelompok dari *propensity score* kelompok perlakuan dan *propensity score* kelompok kontrol.

$$S_T(x) = (\text{banyaknya } \hat{e}_T(x) \text{ yang teramati } \leq x) / n_T$$

$$S_C(x) = (\text{banyaknya } \hat{e}_C(x) \text{ yang teramati } \leq x) / n_C$$

Maka statistik uji yang digunakan adalah

$$KS = \sup |S_T(x) - S_C(x)| \quad (8)$$

Keputusan:

Tolak  $H_0$  dengan tingkat signifikansi  $\alpha$  jika KS lebih besar daripada kuantil  $(1-\alpha)$  [18].

#### F. Pengguna Napza Suntik (Penasun) dan HIV/AIDS

Narkotika, psikotropika, alkohol dan bahan adiktif lainnya atau biasa disingkat dengan napza. Napza suntik adalah napza yang disuntikkan ke dalam tubuh dengan tujuan untuk mempercepat mendapatkan pengaruhnya atau reaksinya, dilakukan karena alasan ekonomis, tidak banyak yang terbuang bila dibandingkan dengan cara inhalasi/dibakar dan alasan solidaritas kelompok dan gaya hidup [22].

*Acquired Immunodeficiency Syndrome* (AIDS) adalah sekumpulan gejala dan infeksi (sindrom) yang timbul karena rusaknya sistem kekebalan tubuh manusia akibat infeksi virus HIV atau infeksi virus-virus lain yang mirip yang menyerang spesies lainnya (SIV, FIV, dan lain-lain). Virusnya sendiri bernama *Human Immunodeficiency Virus*/HIV yaitu virus yang memperlemah kekebalan pada tubuh manusia [14].

Penderita AIDS di masyarakat digolongkan ke dalam dua kategori yaitu penderita yang mengidap HIV dan telah menunjukkan gejala klinis (penderita AIDS) dan penderita yang mengidap HIV tetapi belum menunjukkan gejala klinis (penderita HIV). Masa inkubasi penyakit ini yaitu mulai terjadinya infeksi sampai timbulnya gejala penyakit sangat lama (sampai 5 tahun atau lebih) dan karena infeksi HIV dianggap seumur hidup maka resiko terjadinya penyakit akan berlanjut selama hidup pengidap virus HIV [23].

Faktor-faktor yang mempengaruhi penasun menderita AIDS, yaitu

##### 1. Jenis Kelamin

Penyakit HIV/AIDS dapat menyerang semua orang. Laki-laki lebih banyak terkena HIV/AIDS dibandingkan perempuan. Jenis kelamin berkaitan dengan peran kehidupan dan perilaku berbeda antara laki-laki dan perempuan dalam masyarakat. Dalam hal menjaga kesehatan biasanya perempuan lebih memperhatikan kesehatannya dibandingkan dengan laki-laki. Perbedaan pola perilaku sakit juga dipengaruhi oleh jenis kelamin, perempuan lebih sering mengobati dirinya dibandingkan laki-laki [24].

##### 2. Pendidikan

Menurut UU No 12 tahun 2012, pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa, dan negara. Salah satu cara yang dapat dilakukan untuk mengendalikan diri agar tidak terkena suatu penyakit adalah mencegah terkena suatu penyakit pada diri sendiri bahkan menularkan pada orang lain, karena semakin tinggi pendidikan seseorang maka semakin banyak pengetahuan yang dia miliki sehingga semakin tinggi dan baik upaya pencegahan yang dilakukan oleh seseorang untuk menghindari perilaku yang dapat menyebabkan terjangkitnya suatu penyakit [24].

##### 3. Pasangan Tetap

Pasangan tetap diartikan sebagai sebuah hubungan yang berkelanjutan dalam jangka waktu yang relatif panjang disertai dengan aktifitas seksual dan seringkali melibatkan komitmen pribadi terhadap hubungan tersebut. Seberapa penting dan

seberapa jauh aktivitas seksual dengan pasangan tetap ini sangat bervariasi, bergantung pada sifat hubungan tersebut [25].

#### 4. Pasangan Tidak Tetap

Pasangan tidak tetap mencakup pasangan hubungan seksual dalam satu waktu tertentu atau berulang tetapi tidak ada keterlibatan emosional atau menuntut suatu pengharapan tertentu atas hubungan yang dilakukannya. Pasangan semacam ini ditemui dalam sejumlah konteks sosial dan dalam berbagai kategori pasangan. Hubungan ini mencakup hubungan sebagai teman tetapi memiliki ketertarikan untuk berhubungan seks atau seseorang yang dikenal pada suatu lingkungan sosial tertentu (misalnya mall, kampus, saat membeli obat atau di jalanan) dan sepakat untuk melakukan hubungan seksual. Wanita yang bersedia untuk menjalin hubungan semacam ini biasa disebut dengan *perek* atau perempuan ekperimental dan melibatkan perilaku berisiko lainnya (misalnya pemakaian napza suntik, hubungan seks tanpa kondom) [24].

#### 5. Pemakaian Kondom

Perilaku seks berisiko lain yang dilakukan oleh penasun, yaitu membeli jasa seks tanpa menggunakan kondom. Hampir separuh penasun kena HIV, maka dengan perilaku seks berisiko tersebut akan memperluas penularan selanjutnya. Hanya dengan menghindari penggunaan bersama alat suntik yang tidak steril serta penggunaan kondom pada setiap kegiatan seks yang akan mencegah penyebaran HIV yang lebih luas tidak hanya pada sesama pengguna napza suntik, tetapi juga kelompok lain yaitu kelompok perilaku seks berisiko [26].

Pemakaian kondom pada kalangan penasun sangat tidak konsisten dan bersifat problematis padahal mereka umumnya memiliki banyak pasangan seks yang berisiko tinggi maupun rendah. Terdapat tiga pola pemakaian kondom, yaitu mereka yang menggunakan kondom secara konsisten, mereka yang menggunakan kondom secara tidak konsisten, termasuk yang baru mulai memakai kondom dan mereka yang tidak pernah menggunakan kondom sama sekali [25].

#### 6. Penggunaan Jarum Suntik

Meningkat tajamnya prevalensi HIV pada penasun disebabkan oleh penggunaan jarum dan alat suntik yang tidak steril ditambah dengan praktek penyuntikan berkelompok. Penelitian di beberapa negara mendapatkan perilaku kelompok ini sangat rentan tertular HIV dan penyakit lain melalui penggunaan jarum suntik secara bergantian tanpa melakukan sterilisasi yang memadai [1].

Jaringan pemakaian napza suntik merupakan titik kunci dari perilaku berisiko, hubungan seksual berisiko dan sumber infeksi HIV di kalangan penasun. Jaringan ini mengindikasikan sampai batas mana penyuntikan napza berisiko dan potensi pajanan HIV bertemu dengan hubungan seksual berisiko tinggi dan mengarah ke potensi pajanan yang melintasi wilayah dan kelompok sosial. Besarnya jaringan penyuntikan, stabilitas dan tingkat pertukaran pasangan, pasangan penasun yang kemudian menjadi pasangan penasun lain dari berbagai daerah dan karakteristik demografis serta frekuensi praktek penyuntikan berisiko akan membentuk suatu profil risiko jaringan penyuntikan [25].

### III. METODOLOGI PENELITIAN

#### A. Metode Penelitian

Adapun metode yang digunakan untuk menjawab rumusan masalah pada penelitian ini adalah:

1. Mendapatkan estimasi *propensity score* dengan menggunakan regresi logistik

a. Diberikan distribusi probabilitas untuk setiap pasangan

$$(x_i, z_i), \text{ adalah } f(x_i) = e(x_i)^{z_i} (1 - e(x_i))^{n - z_i},$$

dengan

$$e(x_i) = \frac{\exp\left(\sum_{j=0}^k \beta_j x_{ij}\right)}{1 + \exp\left(\sum_{j=0}^k \beta_j x_{ij}\right)},$$

atau dapat ditulis dalam bentuk matriks sebagai berikut

$$e(\mathbf{x}_i) = \frac{\exp(\boldsymbol{\beta}'\mathbf{x}_i)}{[1 + \exp(\boldsymbol{\beta}'\mathbf{x}_i)]},$$

b. Membentuk fungsi *likelihood* merupakan fungsi distribusi probabilitas bersama, dapat ditulis sebagai berikut

$$l(\boldsymbol{\beta}) = \prod_{i=1}^n e(\mathbf{x}_i)^{z_i} (1 - e(\mathbf{x}_i))^{n - z_i}.$$

c. Memaksimumkan  $\ln l(\boldsymbol{\beta})$  atau disebut juga  $\ln$  *likelihood* yang dinotasikan sebagai

$$L(\boldsymbol{\beta}) = \ln[l(\boldsymbol{\beta})].$$

d. Mendiferensialkan  $L(\boldsymbol{\beta})$  terhadap  $\boldsymbol{\beta}$  dan menyamakannya dengan nol.

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_a} = 0.$$

e. Nilai  $\boldsymbol{\beta}$  diestimasi dengan metode numerik karena persamaannya bersifat nonlinier sehingga metode yang digunakan adalah metode iterasi Newton-Raphson dengan rumus:

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - \left(\mathbf{H}^{(t)}\right)^{-1} \mathbf{q}; t = 1, 2, \dots \text{sampai konvergen}$$

dengan

$$\mathbf{q} = \left[ \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_0}, \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_1}, \dots, \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_k} \right]$$

$\mathbf{H}$  matriks Hessian dengan elemen-elemen  $h_{ab} = \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b}$

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1k} \\ h_{21} & h_{22} & \dots & h_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k1} & h_{k2} & \dots & h_{kk} \end{pmatrix}.$$

2. Aplikasi metode *propensity score* digunakan pada data penderita HIV/AIDS dengan langkah-langkah sebagai berikut:

a. Statistik deskriptif pada data dengan membuat grafik dan tabulasi data berdasarkan variabel.

- Menentukan variabel *confounding*, selanjutnya variabel *confounding* dinotasikan  $z$  dengan parameter  $\tau$ .
- Menghitung nilai estimasi *propensity score* yang telah diperoleh pada langkah (1).
- Melakukan pembobotan pada *propensity score*.
- Balancing* dengan cara menguji apakah *propensity score* dari kelompok perlakuan dan kontrol memiliki distribusi yang sama pada setiap kovariat.
- Pengujian signifikansi dan interpretasi variabel *confounding*.

## B. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder berupa data kasus penderita HIV/AIDS yang telah disurvei oleh LSM ORBIT Surabaya tahun 2013.

## C. Variabel Penelitian

Variabel respon dan variabel prediktor yang digunakan dalam penelitian ini adalah

Tabel 1.  
Variabel Penelitian

Variabel	Deskripsi	Kategori
$y$	Status HIV/AIDS	0 =Negatif 1 =Positif
Variabel Identitas Klien		
$x_1$	Jenis Kelamin	1 = Laki-Laki 2 = Perempuan
$x_2$	Usia	-
$x_3$	Pendidikan Terakhir	1 = Tidak Bersekolah 2 = SMP 3 = SMA 4 = PT 5 = Tidak Menjawab
$x_4$	Penghasilan	1 = < 500.000 2 = 500.000-1.000.000 3 = > 1.000.000 4 = Tidak Menjawab
Pola Perilaku		
$x_5$	Status Nikah	1 = Kawin 2 = Cerai 3 = Tidak Kawin 4 = Tidak Menjawab
$x_6$	Pasangan Tetap	1 = Ada, Laki-Laki 2 = Ada, Perempuan 3 = Tidak Ada
$x_7$	Pasangan Tidak Tetap	1 = Ada, Laki-Laki 2 = Ada, Perempuan 3 = Ada, Laki-Laki dan Perempuan 4 = Tidak Ada
$x_8$	Selalu Pakai Kondom	1 = Ya 2 = Tidak
Riwayat Penggunaan Jarum Suntik		
$x_9$	Zat Yang Disuntikkan	1 = Putau 2 = Buphre 3 = Anti Depresan 4 = Campuran

Lanjutan Tabel 1.

Variabel	Deskripsi	Kategori
$x_{10}$	Frekuensi Suntik	1 = 1-3 kali/hari 2 = 1-3 kali/minggu 3 = 1-3 kali/bulan
$x_{11}$	Pernah Berbagi Jarum Suntik	1 = Ya, Disterilkan 2 = Ya, Tidak Disterilkan 3 = Tidak Pernah
$x_{12}$	Selalu Pakai Jarum Steril	1 = Ya 2 = Tidak
$x_{13}$	Selalu Pakai Jarum Untuk Sendiri	1 = Ya 2 = Tidak

## IV. HASIL DAN PEMBAHASAN

### A. Estimasi *Propensity Score*

Estimasi *propensity score* menggunakan regresi logistik dengan metode MLE yaitu dengan cara estimasi parameter pada model. Diberikan distribusi probabilitas untuk setiap pasangan  $(x_i, z_i)$  dimana  $z_i$  merupakan variabel *confounding*, adalah

$$f(x_i) = e(x_i)^{z_i} (1 - e(x_i))^{n - z_i},$$

hubungan antara  $e(x_i)$  dan  $x_i$  dapat dijelaskan oleh fungsi

logistik pada persamaan berikut

$$e(x_i) = \frac{1}{1 + \exp \left[ - \left( \beta_0 + \sum_{j=1}^k \beta_j x_{ij} \right) \right]}, \quad i = 1, 2, \dots, n$$

$$e(x_i) = \frac{1}{1 + \exp \left[ - \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]}$$

$$e(x_i) = \frac{1}{1 + \exp \left[ - \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]} \frac{\exp \left[ \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]}{\exp \left[ \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]}$$

$$e(x_i) = \frac{\exp \left[ \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]}{1 + \exp \left[ \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]}$$

Dalam bentuk matriks dapat ditulis sebagai berikut,

$$e(\mathbf{x}_i) = \frac{\exp[(\boldsymbol{\beta}'\mathbf{x}_i)]}{1 + \exp[(\boldsymbol{\beta}'\mathbf{x}_i)]},$$

dengan

$$\boldsymbol{\beta} = (\beta_0 \quad \beta_1 \quad \beta_2 \quad \cdots \quad \beta_k), \quad \mathbf{x}_i = \begin{pmatrix} 1 \\ x_{i1} \\ x_{i2} \\ \vdots \\ x_{ik} \end{pmatrix}.$$

Antar pengamatan diasumsikan independen, sehingga fungsi *likelihood* dapat diperoleh sebagai berikut

$$l(\boldsymbol{\beta}) = \prod_{i=1}^n e(\mathbf{x}_i)^{z_i} (1 - e(\mathbf{x}_i))^{n - z_i}$$

$$l(\boldsymbol{\beta}) = \prod_{i=1}^n e(\mathbf{x}_i)^{z_i} (1 - e(\mathbf{x}_i))^{-z_i} (1 - e(\mathbf{x}_i))^n$$

$$l(\boldsymbol{\beta}) = \prod_{i=1}^n \left( \frac{e(\mathbf{x}_i)}{(1 - e(\mathbf{x}_i))} \right)^{z_i} (1 - e(\mathbf{x}_i))^n$$

$$l(\boldsymbol{\beta}) = \left\{ \prod_{i=1}^n \exp \left[ \log \left( \frac{e(\mathbf{x}_i)}{(1 - e(\mathbf{x}_i))} \right)^{z_i} \right] \right\} \left\{ \prod_{i=1}^n (1 - e(\mathbf{x}_i))^n \right\}$$

Logit ke- $i$  adalah  $\exp \left[ \sum_{i=1}^n z_i \boldsymbol{\beta}' \mathbf{x}_i \right]$  dan bentuk

$[1 - e(\mathbf{x}_i)] = [1 + \exp(\boldsymbol{\beta}' \mathbf{x}_i)]^{-1}$ , diperoleh *likelihood* sebagai berikut

$$l(\boldsymbol{\beta}) = \left\{ \exp \left[ \sum_{i=1}^n z_i \boldsymbol{\beta}' \mathbf{x}_i \right] \right\} \left\{ \prod_{i=1}^n n [1 + \exp(\boldsymbol{\beta}' \mathbf{x}_i)]^{-1} \right\},$$

Sehingga diperoleh fungsi log *likelihood*

$$L(\boldsymbol{\beta}) = \left( \sum_{i=1}^n z_i \boldsymbol{\beta}' \mathbf{x}_i \right) - \sum_{i=1}^n n \log [1 + \exp(\boldsymbol{\beta}' \mathbf{x}_i)].$$

Karena  $z_i$  bersifat biner yaitu  $z_i = 1$  untuk kelompok perlakuan dan  $z_i = 0$  kelompok kontrol, sehingga proses estimasi dibagi menjadi

- Untuk kelompok perlakuan  $z_i = 1$

$$L(\boldsymbol{\beta}) = \sum_{i=1}^n \boldsymbol{\beta}' \mathbf{x}_i - \sum_{i=1}^n n \log [1 + \exp(\boldsymbol{\beta}' \mathbf{x}_i)].$$

Langkah selanjutnya, yaitu memaksimumkan log *likelihood* dengan cara mendifferensialkan  $L(\boldsymbol{\beta})$  terhadap  $\beta_j$  dan menyamakannya dengan nol.

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_a} = 0$$

$$\sum_{i=1}^n \mathbf{x}_{ia} - \sum_{i=1}^n n x_{ia} \left[ \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}{1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)} \right] = 0$$

$$\sum_{i=1}^n \mathbf{x}_{ia} - \sum_{i=1}^n n \hat{e}(\mathbf{x}_i) x_{ia} = 0, \quad a = j = 0, 1, 2, \dots, k.$$

Dengan  $\hat{e}(\mathbf{x}_i) = \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}{1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}$  menyatakan estimasi  $e(\mathbf{x}_i)$  dari

dengan menggunakan MLE. Karena hasil differensial pertama tidak *close form*, nilai  $\beta_j$  diestimasi dengan iterasi Newton Raphson. Estimasi varian kovarian dikembangkan menurut teori MLE yang menyatakan bahwa estimasi varian kovarian diperoleh dari turunan kedua fungsi log *likelihood*.

Turunan kedua dari fungsi log *likelihood* yaitu:

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a^2} = - \sum_{i=1}^n n x_{ia}^2 \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}{[1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)]^2} = - \sum_{i=1}^n n x_{ia}^2 \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i))$$

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b} = - \sum_{i=1}^n x_{ia} x_{ib} n \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}{[1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)]^2} = - \sum_{i=1}^n x_{ia} x_{ib} n \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i))$$

$a, b = 0, 1, 2, \dots, k$ .

Sehingga diperoleh matriks varian kovarian dari estimasi parameter melalui invers matriks,

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \left\{ \mathbf{X}' \text{Diag} [n \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i))] \mathbf{X} \right\}^{-1},$$

dengan

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$\text{Diag} [n \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i))]$  merupakan  $n \times n$  matriks diagonal

dengan elemen diagonal utama  $[\hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i))]$ . Akar kuadrat dari elemen-elemen diagonal utama adalah estimasi standar *error* dari estimasi parameter model. Dimana untuk mendapatkan nilai estimasi  $\boldsymbol{\beta}$  dari penyelesaian turunan pertama fungsi log *likelihood* yang nonlinier digunakan metode iterasi Newton-Raphson dengan rumus

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - (\mathbf{H}^{(t)})^{-1} \mathbf{q}^{(t)}; t = 1, 2, \dots \text{ sampai konvergen}$$

dengan

$\mathbf{H}$  matriks Hessian sebagai berikut:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k1} & h_{k2} & \cdots & h_{kk} \end{bmatrix}; h_{ab} = \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b}.$$

$$\mathbf{q}' = \left( \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_0}, \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_1}, \dots, \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_k} \right).$$

Untuk setiap langkah iterasi ke- $t$ , berlaku

$$h_{ab}^{(t)} = \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b} \Big|_{\beta^{(t)}} = - \sum_{i=1}^n x_{ia} x_{ib} n \hat{e}(\mathbf{x}_i)^{(t)} (1 - \hat{e}(\mathbf{x}_i)^{(t)})$$

$$q_j^{(t)} = \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_j} \Big|_{\beta^{(t)}} = \sum_{i=1}^n (1 - n \hat{e}(\mathbf{x}_i)^{(t)}) x_{ij}$$

$$\hat{e}(\mathbf{x}_i)^{(t)} = \frac{\exp(\hat{\boldsymbol{\beta}}^{(t)} \mathbf{x}_i)}{1 + \exp(\hat{\boldsymbol{\beta}}^{(t)} \mathbf{x}_i)}$$

Sehingga diperoleh

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} + \left\{ \mathbf{X}' \text{Diag} \left[ n \hat{e}(\mathbf{x}_i)^{(t)} (1 - \hat{e}(\mathbf{x}_i)^{(t)}) \right] \mathbf{X} \right\}^{-1} \mathbf{X}' (1 - \mathbf{m}^{(t)})$$

dengan  $\mathbf{m}^{(t)} = n \hat{e}(\mathbf{x}_i)^{(t)}$ .

Setelah didapat  $\mathbf{H}$  dan  $\mathbf{q}$  maka dilakukan iterasi dengan langkah-langkah sebagai berikut:

1. Menentukan *starting value* atau menentukan nilai awal dari  $\hat{\boldsymbol{\beta}}$  pada saat iterasi pertama yaitu  $\hat{\boldsymbol{\beta}} = \mathbf{0}$ .
2. Mulai dari iterasi pertama atau  $t = 0$  kemudian dilakukan iterasi dengan menghitung  $\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - (\mathbf{H}^{(t)})^{-1} \mathbf{q}^{(t)}$ .
3. Jika  $\|\boldsymbol{\beta}^{(t+1)} - \boldsymbol{\beta}^{(t)}\| \leq \Theta$  dimana  $\Theta$  adalah bilangan yang sangat kecil maka iterasi berhenti dan didapatkan hasil perhitungan, jika tidak maka ulangi langkah sebelumnya.

- Untuk kelompok kontrol  $z_i = 0$

$$L(\boldsymbol{\beta}) = - \sum_{i=1}^n n \log [1 + \exp(\boldsymbol{\beta}' \mathbf{x}_i)].$$

Langkah selanjutnya, yaitu memaksimumkan log *likelihood* dengan cara mendifferensialkan  $L(\boldsymbol{\beta})$  terhadap  $\beta_j$  dan menyamakannya dengan nol.

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_a} = 0$$

$$- \sum_{i=1}^n n x_{ia} \left[ \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}{1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)} \right] = 0$$

$$- \sum_{i=1}^n n \hat{e}(\mathbf{x}_i) x_{ia} = 0, \quad a = j = 0, 1, 2, \dots, k.$$

Dengan  $\hat{e}(\mathbf{x}_i) = \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}{1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}$  menyatakan estimasi  $e(\mathbf{x}_i)$  dari

dengan menggunakan MLE. Karena hasil differensial pertama tidak *close form*, nilai  $\beta_j$  diestimasi dengan iterasi Newton Raphson. Estimasi varian kovarian dikembangkan menurut teori MLE yang menyatakan bahwa estimasi varian kovarian diperoleh dari turunan kedua fungsi log *likelihood*.

Turunan kedua dari fungsi log *likelihood* yaitu:

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a^2} = - \sum_{i=1}^n n x_{ia}^2 \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}{[1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)]^2} = - \sum_{i=1}^n n x_{ia}^2 \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i))$$

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b} = - \sum_{i=1}^n x_{ia} x_{ib} n \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)}{[1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i)]^2} = - \sum_{i=1}^n x_{ia} x_{ib} n \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i))$$

$$a, b = 0, 1, 2, \dots, k.$$

Sehingga diperoleh matriks varian kovarian dari estimasi parameter melalui invers matriks,

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \left\{ \mathbf{X}' \text{Diag} \left[ n \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i)) \right] \mathbf{X} \right\}^{-1},$$

dengan

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$\text{Diag} \left[ n \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i)) \right]$  merupakan  $n \times n$  matriks diagonal dengan elemen diagonal utama  $\left[ \hat{e}(\mathbf{x}_i) (1 - \hat{e}(\mathbf{x}_i)) \right]$ . Akar kuadrat dari elemen-elemen diagonal utama adalah estimasi standar *error* dari estimasi parameter model. Dimana untuk mendapatkan nilai estimasi  $\boldsymbol{\beta}$  dari penyelesaian turunan pertama fungsi log *likelihood* yang nonlinier digunakan metode iterasi Newton-Raphson dengan rumus

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - (\mathbf{H}^{(t)})^{-1} \mathbf{q}^{(t)}; t = 1, 2, \dots \text{ sampai konvergen}$$

dengan

$\mathbf{H}$  matriks Hessian sebagai berikut:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1k} \\ h_{21} & h_{22} & \dots & h_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k1} & h_{k2} & \dots & h_{kk} \end{bmatrix}; h_{ab} = \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b}$$

$$\mathbf{q}' = \left( \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_0}, \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_1}, \dots, \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_k} \right)$$

Untuk setiap langkah iterasi ke- $t$ , berlaku

$$h_{ab}^{(t)} = \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b} \Big|_{\boldsymbol{\beta}^{(t)}} = - \sum_{i=1}^n x_{ia} x_{ib} n \hat{e}(\mathbf{x}_i)^{(t)} (1 - \hat{e}(\mathbf{x}_i)^{(t)})$$

$$q_j^{(t)} = \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_j} \Big|_{\boldsymbol{\beta}^{(t)}} = - \sum_{i=1}^n n \hat{e}(\mathbf{x}_i)^{(t)} x_{ij}$$

$$\hat{e}(\mathbf{x}_i)^{(t)} = \frac{\exp(\hat{\boldsymbol{\beta}}^{(t)} \mathbf{x}_i)}{1 + \exp(\hat{\boldsymbol{\beta}}^{(t)} \mathbf{x}_i)}$$

Sehingga diperoleh

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} + \left\{ \mathbf{X}' \text{Diag} \left[ n \hat{e}(\mathbf{x}_i)^{(t)} (1 - \hat{e}(\mathbf{x}_i)^{(t)}) \right] \mathbf{X} \right\}^{-1} (-\mathbf{X}' \mathbf{m}^{(t)})$$

dengan  $\mathbf{m}^{(t)} = n \hat{e}(\mathbf{x}_i)^{(t)}$ .

Setelah didapat  $\mathbf{H}$  dan  $\mathbf{q}$  maka dilakukan iterasi dengan langkah-langkah sebagai berikut:

1. Menentukan *starting value* atau menentukan nilai awal dari  $\hat{\boldsymbol{\beta}}$  pada saat iterasi pertama yaitu  $\hat{\boldsymbol{\beta}} = \mathbf{0}$ .
2. Mulai dari iterasi pertama atau  $t = 0$  kemudian dilakukan iterasi dengan menghitung  $\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - (\mathbf{H}^{(t)})^{-1} \mathbf{q}^{(t)}$ .
3. Jika  $\|\boldsymbol{\beta}^{(t+1)} - \boldsymbol{\beta}^{(t)}\| \leq \Theta$  dimana  $\Theta$  adalah bilangan yang sangat kecil maka iterasi berhenti dan didapatkan hasil perhitungan, jika tidak maka ulangi langkah sebelumnya.

## B. Analisis Propensity Score

### 1. Deskriptif Data

Berdasarkan Tabel 2, penasun memiliki kesadaran yang rendah untuk menggunakan kondom dimana penasun yang memiliki kesadaran untuk menggunakan kondom hanya 17 orang dari 218 orang. Selain itu, penasun dengan status positif HIV/AIDS yang tidak selalu menggunakan kondom memiliki

persentase yang lebih besar yaitu 87,5% dibandingkan selalu menggunakan kondom dengan persentase sebesar 12,5%.

Tabel 2.

Tabulasi Silang Status HIV/AIDS dan Selalu Pakai Kondom

Status HIV	Selalu Pakai Kondom		Total
	Ya	Tidak	
Negatif	11 (6,5%)	159 (93,5%)	170 (100%)
Positif	6 (12,5%)	42 (87,5%)	48 (100%)
Total	17 (7,8%)	201 (92,2%)	218 (100%)

Sumber: Diolah menggunakan SPSS

### 2. Pemilihan Variabel *Confounding*

Adanya hubungan masing-masing variabel yang termasuk pola perilaku dengan variabel yang termasuk riwayat penggunaan jarum suntik menunjukkan bahwa variabel yang termasuk pola perilaku tersebut merupakan variabel *confounding*. Setelah melakukan pengujian dengan statistik uji *chi square*, ternyata variabel yang termasuk pola perilaku selalu pakai kondom memiliki hubungan dengan 2 variabel yang termasuk riwayat penggunaan jarum suntik yaitu selalu pakai jarum steril dan selalu pakai jarum untuk sendiri. Sedangkan variabel pola perilaku yaitu status nikah, pasangan tetap dan pasangan tidak tetap hanya memiliki hubungan dengan satu variabel yang termasuk riwayat penggunaan jarum suntik. Sehingga variabel *confounding* yang dipilih adalah variabel selalu pakai kondom.

### 3. Model Propensity Score

$$e(x_i) = \frac{\exp(k)}{1 + \exp(k)}$$

dengan

$$k = -12,573 + 2,334x_1 + 0,047x_2 - 17,383x_{3,1} - 2,063x_{3,2} + 0,098x_{3,3} + 4,274x_{3,4} - 16,205x_{4,1} + 2,961x_{4,2} + 0,550x_{4,3} - 1,122x_{5,1} - 19,526x_{5,2} - 0,734x_{5,3} + 1,135x_{6,1} + 1,159x_{6,2} + 3,345x_{7,1} + 1,880x_{7,2} - 18,762x_{7,3} + 1,846x_{9,1} - 16,381x_{9,2} - 15,921x_{9,3} - 0,588x_{10,1} + 0,301x_{10,2} + 2,149x_{11,1} - 0,690x_{11,2} + 1,312x_{12} + 1,127x_{13}$$

### 4. Evaluasi *Propensity Score*

Hipotesis

$$H_0 : F_T(\hat{e}(x_{ij})) = F_C(\hat{e}(x_{ij})) \text{ untuk semua } \hat{e}(x_{ij})$$

$$H_1 : F_T(\hat{e}(x_{ij})) \neq F_C(\hat{e}(x_{ij})) \text{ untuk paling tidak satu } \hat{e}(x_{ij})$$

Taraf signifikansi:  $\alpha = 5\%$ .

Statistik uji yang digunakan: Jika  $S_T(x_{ij})$  dan  $S_C(x_{ij})$  berturut-turut adalah fungsi-fungsi distribusi kelompok dari *propensity score* kelompok perlakuan dan *propensity score* kelompok kontrol.



$$S_T(x_{ij}) = \left( \text{banyaknya } \hat{e}_T(x_{ij}) \text{ yang teramati } \leq x_{ij} \right) / n_T$$

$$S_C(x_{ij}) = \left( \text{banyaknya } \hat{e}_C(x_{ij}) \text{ yang teramati } \leq x_{ij} \right) / n_C$$

Maka statistik uji yang digunakan adalah

$$KS = \max \left| S_T(x_{ij}) - S_C(x_{ij}) \right|$$

Daerah kritis: Tolak  $H_0$  dengan tingkat signifikansi  $\alpha$  jika

$$KS_{hitung} > KS_{tabel} \text{ dimana } KS_{tabel} = 0,3435$$

Keputusan:

Tabel 3.  
Balancing Kovariat

Variabel	KS <sub>hitung</sub>	Keputusan
$x_1$	0,00	Gagal tolak $H_0$
$x_2$	0,14	Gagal tolak $H_0$
$x_{3(1)}$	0,00	Gagal tolak $H_0$
$x_{3(2)}$	0,01	Gagal tolak $H_0$
$x_{3(3)}$	0,04	Gagal tolak $H_0$
$x_{3(4)}$	0,04	Gagal tolak $H_0$
$x_{4(1)}$	0,00	Gagal tolak $H_0$
$x_{4(2)}$	0,07	Gagal tolak $H_0$
$x_{4(2)}$	0,11	Gagal tolak $H_0$
$x_{5(1)}$	0,03	Gagal tolak $H_0$
$x_{5(2)}$	0,00	Gagal tolak $H_0$
$x_{5(3)}$	0,01	Gagal tolak $H_0$
$x_{6(1)}$	0,02	Gagal tolak $H_0$
$x_{6(2)}$	0,00	Gagal tolak $H_0$
$x_{7(1)}$	0,00	Gagal tolak $H_0$
$x_{7(2)}$	0,02	Gagal tolak $H_0$
$x_{7(3)}$	0,00	Gagal tolak $H_0$
$x_{9(1)}$	0,09	Gagal tolak $H_0$
$x_{9(2)}$	0,00	Gagal tolak $H_0$
$x_{9(3)}$	0,00	Gagal tolak $H_0$
$x_{10(1)}$	0,05	Gagal tolak $H_0$
$x_{10(2)}$	0,01	Gagal tolak $H_0$
$x_{11(1)}$	0,09	Gagal tolak $H_0$
$x_{11(2)}$	0,09	Gagal tolak $H_0$
$x_{12(1)}$	0,01	Gagal tolak $H_0$
$x_{13(1)}$	0,04	Gagal tolak $H_0$

Kesimpulan: tidak terdapat perbedaan antara kelompok perlakuan dan kelompok kontrol (*balance*).

## 5. Pengujian Signifikansi Variabel *Confounding* Selalu Pakai Kondom

Hipotesis

$H_0: \tau = 0$  (Variabel *confounding* selalu pakai kondom tidak signifikan)

$H_1: \tau \neq 0$  (Variabel *confounding* selalu pakai kondom signifikan)

Taraf signifikansi:  $\alpha = 10\%$

Statistik uji yang digunakan:

$$W = \frac{\hat{\tau}}{Se(\hat{\tau})}$$

Daerah kritis:

Tolak  $H_0$  jika  $|W| > Z_{\alpha/2}$  atau  $p\text{-value} < \alpha$

Keputusan:

Tolak  $H_0$  karena  $p\text{-value} < \alpha$  ( $0,07772 < 0,10$ )

Tabel 4. Pengujian Signifikansi Variabel *Confounding*

Koefisien	Estimasi	Standar Error	Wald	P-value
Konstanta	0,1320	0,0452	2,920	0,00387
$x_8$	0,2210	0,1247	1,773	0,07772

Sumber: Diolah menggunakan *package R*

Kesimpulan: Variabel *confounding* selalu pakai kondom signifikan.

Interpretasi:

Berdasarkan Tabel 4.16, Penasun yang selalu menggunakan kondom dapat terkena HIV/AIDS 1,25 kali dibandingkan dengan tidak selalu menggunakan kondom.

## V. PENUTUP

### A. Kesimpulan

Berdasarkan hasil analisis data dan pembahasan, maka kesimpulan yang diperoleh:

1. Estimasi *propensity score* menggunakan regresi logistik dengan metode MLE, pada turunan pertama fungsi log *likelihood* terhadap diperoleh:

a. Untuk kelompok perlakuan

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_a} = \sum_{i=1}^n \mathbf{x}_{ia} - \sum_{i=1}^n n_i \hat{e}(\mathbf{x}_i) \mathbf{x}_{ia} = 0,$$

karena hasil tidak *close form*, maka dilakukan iterasi Newton-Raphson dengan rumus:

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} + \left\{ \mathbf{x}' \text{Diag} \left[ n \hat{e}(\mathbf{x}_i)^{(t)} \left( 1 - \hat{e}(\mathbf{x}_i)^{(t)} \right) \right] \mathbf{x} \right\}^{-1} \mathbf{x}' (1 - \mathbf{m}^{(t)})$$

dengan  $\mathbf{m}^{(t)} = n \hat{e}(\mathbf{x}_i)^{(t)}$ .

Jika  $\|\beta^{(t+1)} - \beta^{(t)}\| \leq \Theta$  dimana  $\Theta$  adalah bilangan yang sangat kecil maka iterasi berhenti dan didapatkan hasil perhitungan.

b. Untuk kelompok kontrol

$$\frac{\partial L(\beta)}{\partial \beta_a} = -\sum_{i=1}^n n_i \hat{e}(x_i) x_{ia} = 0,$$

karena hasil tidak *close form*, maka dilakukan iterasi Newton-Raphson dengan rumus:

$$\beta^{(t+1)} = \beta^{(t)} + \left\{ \mathbf{x}' \text{Diag} \left[ n \hat{e}(x_i)^{(t)} (1 - \hat{e}(x_i)^{(t)}) \right] \mathbf{x} \right\}^{-1} \begin{pmatrix} -\mathbf{x}' \mathbf{m}^{(t)} \end{pmatrix}$$

dengan  $\mathbf{m}^{(t)} = n \mathbf{e}(x_i)^{(t)}$ .

Jika  $\|\beta^{(t+1)} - \beta^{(t)}\| \leq \Theta$  dimana  $\Theta$  adalah bilangan yang sangat kecil maka iterasi berhenti dan didapatkan hasil perhitungan.

2. Pemasun yang selalu menggunakan kondom dapat terkena HIV/AIDS 1,25 kali dibandingkan dengan tidak selalu menggunakan kondom. Hal ini tidak sesuai dengan realita bahwa tingkat pemakaian kondom yang rendah menempatkan posisi berisiko tinggi penularan infeksi HIV/AIDS. Karena pemasun dapat berisiko tinggi terinfeksi HIV/AIDS berawal dari perilaku mereka dalam penggunaan jarum suntik lalu ketika pemasun tersebut terinfeksi HIV/AIDS melalui jarum suntik maka hal tersebut dapat menularkannya kepada pasangan tetap ataupun pasangan tidak tetapnya melalui aktivitas seksual.

## B. Saran

Walaupun hasil penelitian menyebutkan bahwa pemasun sebaiknya menggunakan kondom agar tidak menularkan HIV/AIDS. Namun, langkah lebih baiknya jika pemasun tidak melakukan hubungan seks yang tidak aman dan berhenti menggunakan napza dengan melakukan rehabilitasi terhadap dirinya dan berhenti untuk konsumsi napza.

Berdasarkan hasil penelitian, beberapa hal yang dapat disarankan untuk penelitian selanjutnya, yaitu:

1. Pemilihan variabel *confounding* lebih dari satu sehingga dapat melihat pengaruh perlakuan pada variabel *confounding* lainnya yaitu dengan menggunakan Anacova.
2. *Propensity score* menggunakan model lain seperti regresi linier, CART, MARS dan lain-lain.
3. Metode *propensity score* dapat menggunakan stratifikasi, fungsi kernel, *neural network* dan lain-lain.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Direktorat Jenderal Pendidikan Tinggi (Ditjen Dikti) yang telah memberikan dukungan finansial melalui Beasiswa Pendidikan Pascasarjana Dalam Negeri tahun 2013-2015. Penulis juga menyampaikan ucapan terima kasih kepada pihak Pascasarjana Institut Teknologi Sepuluh Nopember (ITS) Surabaya.

## DAFTAR PUSTAKA

- [1] Besral, B. Utomo, and A. P. Zani, "Potensi Penyebaran HIV dari Pengguna NAPZA Suntik ke Masyarakat Umum," vol. 8, no. 2, pp. 53–58, 2004.
- [2] P. Rosenbaum and D. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, vol. 70, no. 1, pp. 41–55, 1983.
- [3] S. Littnerova, J. Jarkovsky, J. Parenica, T. Pavlik, J. Spinar, and L. Dusek, "Why to Use Propensity Score in Observational Studies? Case Study Based on Data from the Czech Clinical Database AHEAD 2006–09," *Cor Vasa*, vol. 55, no. 4, pp. 383–390, Aug. 2013.
- [4] G. W. Imbens, "The Role of the Propensity Score in Estimating Dose-Response Functions," vol. 87, no. 3, pp. 706–710, 2007.
- [5] D. A. Freedman and R. A. Berk, "Weighting Regressions by Propensity Scores," no. 32, pp. 392–409, 2008.
- [6] P. Rosenbaum and D. B. Rubin, "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score," *J. Am. Stat. Assoc.*, vol. 79, no. No. 387, pp. 516–524, 1984.
- [7] J. K. Lunceford and M. Davidian, "Stratification and Weighting Via the Propensity Score in Estimation of Causal Treatment Effects: A Comparative Study," 2000.
- [8] R. B. D'Agostino, "Tutorial In Biostatistics Propensity Score Method For Bias Reduction In The Comparison Of A Treatment To A Non-Randomized Control Group," vol. 17, pp. 2265–2281, 1998.
- [9] D. F. McCaffrey, G. Ridgeway, and A. R. Moral, "Propensity Score Estimator With Boosted Regression for Evaluating Causal Effect in Observational Studies," *Psychological Method*, vol. Vol. 9, no. No.4, p. 403, 2004.
- [10] S. Setoguchi, S. Schneeweiss, M. A. Brookhart, R. J. Glynn, and E. F. Cook, "NIH Public Access," *Pharmacoepidemiol*, vol. 17, no. 6, pp. 546–555, 2009.
- [11] R. Emsley, M. Lunt, A. Pickles, and G. Dunn, "The Stata Journal," *Stata J.*, vol. Vol. 8, no. No. 3, pp. 334–353, 2008.
- [12] C. D. Newgard, J. R. Hedges, M. Arthur, and R. J. Mullins, "Advanced statistics: the propensity score--a method for estimating treatment effect in observational research," *Acad. Emerg. Med.*, vol. 11, no. 9, pp. 953–61, Sep. 2004.
- [13] I. Praptoraharjo, W. W. Wiebel, O. Kamil, and A. P. Iii, "Jaringan Seksual dan Perilaku Berisiko Pengguna Napza Suntik: Episode Lain Penyebaran HIV di Indonesia," *Ber. Kedokt. Masy.*, vol. 23, no. 3, pp. 106–118, 2007.
- [14] N. L. Pratiwi and H. Basuki, "HIV-AIDS dan Perilaku Seks Tidak Aman di Indonesia," *Bul. Penelit. Sist. Kesehat.*, vol. 14, no. 4, pp. 346–357, 2011.
- [15] N. K. Arniti, "Faktor-Faktor yang Berhubungan dengan Penerimaan Tes HIV oleh Ibu Hamil di Puskesmas Kota Denpasar," Program Studi Ilmu Kesehatan Masyarakat, Universitas Udayana, Denpasar, 2014.
- [16] S. Kristianti, "Dukungan WPS dan Teman Pelanggan terhadap Penggunaan Kondom pada Pelanggan WPS di Semampir Kediri Shinta Kristiantii," *STIKES*, vol. 5, no. 2, 2012.
- [17] I. Yanovitzky, E. Zanutto, and R. Hornik, "Estimating Causal Effects of Public Health Education Campaigns using Propensity Score Methodology," *Eval. Program Plann.*, vol. 28, no. 2, pp. 209–220, May 2005.
- [18] W. W. Daniel, *Statistik Nonparametrik Terapan*. Jakarta: Gramedia, 1978.
- [19] A. Agresti, *Categorical Data Analysis*. New York: John Wiley and Sons, Inc, 1990.
- [20] C. T. Le, *Applied Categorical Data Analysis*. USA: John Wiley and Sons, Inc, 1998.
- [21] D. Hosmer and S. Lemeshow, *Applied Logistic Regression*. USA: John Wiley and Sons, Inc, 1989.
- [22] P. T. Y. S. Suyasa and F. Wijaya, "Resiliensi dan Sikap Terhadap Penyalahgunaan Zat (Studi Pada Remaja)," *J. Psikol.*, vol. 4, no. 2, 2006.
- [23] WHO, "Technical Working Group for The Development of an HIV/AIDS Diagnostic Support Toolkit," p. 2, 2007.
- [24] I. Y. Kumalasari, "Perilaku Berisiko Penyebab Human Immunodeficiency Virus (HIV) Positif (Studi Kasus di Rumah Damai Kelurahan Cepoko Kecamatan Gunungpati Kota Semarang)," Skripsi, Ilmu Kesehatan Masyarakat, Universitas Negeri Semarang, 2013.
- [25] A. Pach, W. Wayne, and I. Praptoraharjo, "Laporan Penelitian Penyebaran HIV di Indonesia: Studi Etnografi tentang Jaringan Seksual dan Perilaku Berisiko Pemakai Napza Suntik (Penasun)," Jakarta Pusat, 2006.
- [26] Komisi Penanggulangan AIDS Nasional, "Ancaman HIV/AIDS di Indonesia Semakin Nyata, Perlu Penanggulangan Lebih Nyata," Jakarta, 2002.