SEMINAR HASIL TESIS

PROPENSITY SCORE Menggunakan MODEL REGRESI LOGISTIK Pada KASUS DATA HIV/AIDS LSM ORBIT SURABAYA

Oleh:
Farida Islamiah (NRP. 1313 201 007)

Dosen Pembimbing:
Dr. Bambang Widjanarko Otok, M.Si

Surabaya, 29 April 2015

$e(x_i)$

JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKONOLOGI SEPULUH NOPEMBER
SURABAYA

PROPENSITY SCORE

Mengurangi bias dalam mengestimasi dampak dari perlakuan pada data yang bersifat observasi.

Rosenbaum dan Rubin (1983)

Studi Observasi → Perlakuan dan Kontrol tidak acak → Bias karena Confounding → Tereduksi

Perlakuan = Kontrol

$e(x_i)$

$e(x_i)$

Kompleks dan sulit diinterpretasikan

Metode Estimasi Propensity Score

D'Agostino (1998) menggunakan analisis diskriminan

McCaffrey (2004) menggunakan model *generalized boosted*

Setouguchi (2008) menggunakan studi simulasi pada *neural network* dan *classification tree*

REGRESI LOGISTIK

$e(x_i)$

Penelitian Tentang Propensity Score

(D'Agostino, 1998) untuk reduksi bias pada perbandingan kelompok perlakuan dan kelompok kontrol dengan menggunakan pencocokan stratifikasi.

Kurth, Walker, Glynn, Chan, Gaziano, Berger dan Robins (2005) membandingkan 5 metode untuk evaluasi jaringan *plasminogen activator* kematian pada 6.269 pasien stroke iskemik yang termasuk dalam daftar pasien stroke di Jerman dengan menggunakan regresi logistik multivariabel, *propensity matching*, regresi *adjustment* dengan *propensity score*, *propensity score* dengan pembobotan dan estimasi pengaruh perlakuan yang dimasukkan dalam populasi.

Gebel dan Vobemer (2014) menggunakan pendekatan *different-in-different propensity score* pada dampak transisi pemerintahan dalam kondisi kesehatan di Jerman.

Regresi Logistik

$e(x_i)$

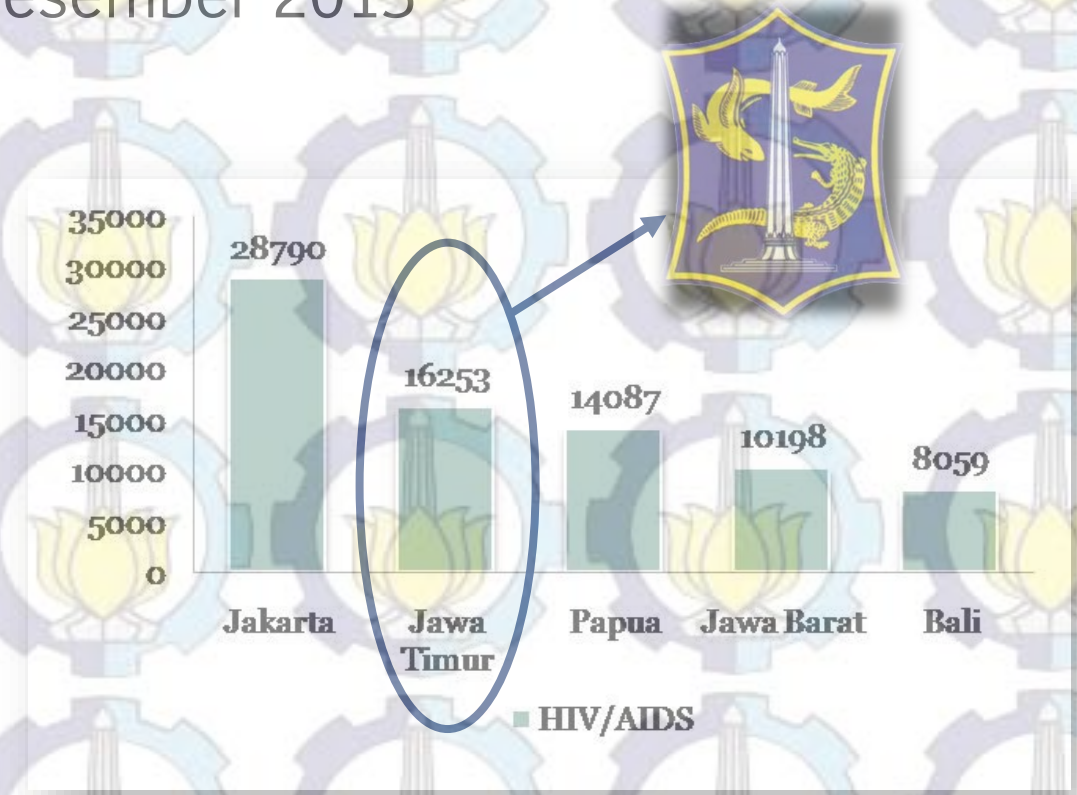
Regresi logistik adalah teknik statistik yang estimasi probabilitas dari suatu peristiwa (variabel respon) berdasarkan faktor-faktor yang diketahui (variabel prediktor) yang diekspektasi mempengaruhi terjadinya peristiwa (perlakuan dalam kasus perhitungan *propensity score*). Variabel respon diasumsikan dua nilai yaitu terjadi atau tidak. Hasil regresi logistik adalah probabilitas (mulai dari 0 hingga 1).

HIV/AIDS

HIV (*Human Immunodefiency Virus*) merupakan virus yang menyerang sistem kekebalan tubuh manusia dan kemudian menyebabkan AIDS. Sedangkan, *Acquired Immuno Defiency Sydrome* (AIDS) merupakan kumpulan gejala penyakit akibatnya sistem kekebalan tubuh yang dirusak.

HIV/AIDS

Berdasarkan laporan situasi perkembangan HIV/AIDS di Indonesia dilaporkan oleh provinsi sampai dengan Desember 2013



$e(x_i)$

Penelitian Regresi Logistik pada Kasus HIV/AIDS

Besral et al. (2004) mengenai besaran potensi penyebaran HIV dari penasun ke masyarakat umum.

Praptoraharjo, Wiebel, Kamil, & Iii (2007) membahas tentang jaringan seksual dan perilaku beresiko penasun dengan analisis data dilakukan yaitu mengembangkan serangkaian kode yang terkait dengan hasil wawancara.

Winarno, Suryoputro dan Shaluhiyah (2008) meneliti tentang faktor-faktor yang berhubungan dengan penggunaan jarum suntik bergantian diantara pengguna napza di kota Semarang.

Pratiwi & Basuki (2011) meneliti tentang hubungan pengetahuan pencegahan HIV/AIDS dan perilaku seks tidak aman pada ramaja usia 15-24 tahun di Indonesia.

Kambu (2012) meneliti faktor-faktor yang mempengaruhi tindakan pencegahan penularan HIV oleh ODHA.

Rumusan Masalah

Bagaimana ?

Tujuan Penelitian

1. Menentukan
2. Mengkaji

Manfaat Penelitian

1. Memberikan wawasan dan literatur
2. Memperoleh informasi

Batasan Penelitian

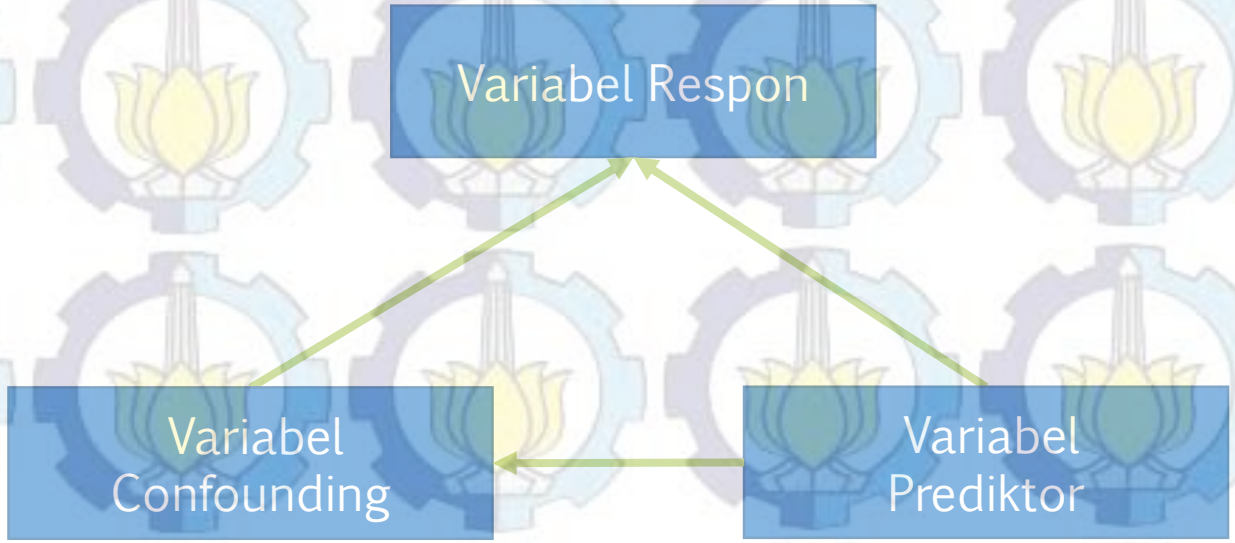
1. Estimasi *propensity score* berdasarkan regresi logistik.
2. Analisis *propensity score* pada kasus penderita HIV/AIDS LSM ORBIT di Surabaya.

1. Data yang digunakan adalah data skunder pada kasus penderita HIV/AIDS di Surabaya berdasarkan kuesioner LSM ORBIT.
2. Metode estimasi *propensity score* yang digunakan adalah *Maximum Likelihood Estimator* (MLE) dan iterasi Newton Raphson.
3. Metode yang digunakan dalam *propensity score* yaitu pembobotan untuk menghilangkan *confounding*. Variabel *confounding* yang digunakan hanya satu.

$e(x_i)$

Variabel Confounding

Proses pemilihan *confounder* dapat berdasarkan teori dan menunjukkan hubungan antara variabel.



H_0: Tidak terdapat hubungan yang signifikan antar variabel

H_1: Terdapat hubungan yang signifikan antar variabel

Taraf signifikansi: α=5%

Statistik uji:

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

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-*value*<α

$e(x_i)$

Propensity score (Rosenbaum dan Rubin, 1983) merupakan probabilitas bersyarat mendapatkan perlakuan tertentu dengan melibatkan kovariat yang diamati.

$$e(x_i) = P(Z_i = 1 | X_i = x_i) \quad (2)$$

Propensity Score

Perlakuan

Kovariat

Regresi logistik digunakan jika variabel respon bersifat kategorik (nominal atau ordinal) dengan variabel-variabel prediktor kontinu maupun kategorik.

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)} \quad (3)$$

Model propensity score menggunakan regresi logistik, variabel respon adalah biner dimana Z_i=1 untuk unit perlakuan dan Z_i=0 untuk unit kontrol dengan model sebagai berikut

$$e(x_i) = P(Z_i = 1 | X_i = x_i) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)} \quad (4)$$

Propensity Score Terbobot

$e(x_i)$

Propensity score terbobot dilakukan ketika estimasi pengaruh perlakuan. Untuk estimasi $E(y_0|z=1)$, misalkan pengamatan ke-i pada sampel pembanding memiliki bobot $w_i=e(x_i)/(1-e(x_i))$, kemungkinan bahwa pengamatan yang dipilih secara random dengan x termasuk dalam perlakuan. Rata-rata bobot y_0 memperoleh perlakuan:

$$\hat{E}\left(y_0|z=1\right)=\frac{\sum_{i\in C} w_i y_i}{\sum_{i\in C} w_i}, \quad (5)$$

Rata-rata bobot y_1 memperoleh perlakuan:

$$\hat{E}\left(y_1|z=1\right)=\sum_{i\in T} y_i / N_T \quad (6)$$

$e(x_i)$

Evaluasi Propensity Score

Evaluasi *propensity score* dilakukan untuk mengetahui apakah kelompok perlakuan dan kontrol berditribusi sama atau *balance*.

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

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

Taraf signifikansi: α=5%

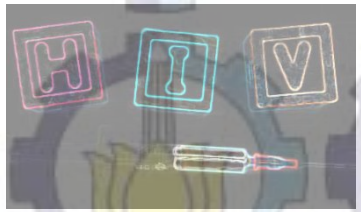
Statistik uji : $$KS = \sup\left|S_T\left(x_{ij}\right) - S_C\left(x_{ij}\right)\right| \qquad (7)$$

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

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

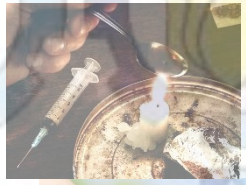
Keputusan: Tolak H0 jika $KS_{hitung} > KS_{tabel}$

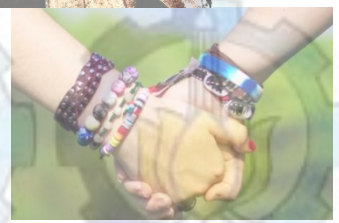
NAPZA Suntik HIV/AIDS



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.

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.





Penasun menghadapi dua risiko untuk terkena HIV/AIDS. Pertama, melalui jarum dan alat suntik yang tercemar yang digunakan secara bersama-sama. Kedua, melalui hubungan seksual terutama bagi mereka yang melakukannya dengan lebih dari satu pasangan atau tanpa menggunakan kondom

Metode Penelitian

Untuk menjawab rumusan masalah pertama

Estimasi propensity score dilakukan dengan maximum likelihood estimation, karena persamaan bersifat nonlinier maka digunakan metode iterasi Newton-Raphson.

1. Membentuk fungsi *likelihood* pada distribusi probabilitas.
2. Memaksimumkan ln $l(\boldsymbol{\beta})$ yang dinotasikan sebagai $L(\boldsymbol{\beta})$.
3. Mendiferensialkan $L(\boldsymbol{\beta})$ terhadap $\boldsymbol{\beta}$ dan menyamakannya dengan nol.
4. Estimasi dibagi dua yaitu kelompok perlakuan dan kelompok kontrol.
5. Pada masing-masing kelompok, nilai $\boldsymbol{\beta}$ diestimasi dengan metode numerik karena persamaannya bersifat nonlinier maka metode yang digunakan adalah metode iterasi Newton-Raphson.

Metode Penelitian

Untuk menjawab rumusan masalah kedua

Aplikasi metode *propensity score* digunakan pada data HIV/AIDS

1. Statistik deskriptif.
2. Menentukan variabel *confounding*, selanjutnya variabel *confounding* dinotasikan z dengan parameter τ.
3. Menghitung nilai estimasi *propensity score* yang telah diperoleh pada jawaban rumusan masalah (1).
4. Melakukan pembobotan pada p*ropensity score*.
5. *Balancing* dengan cara menguji apakah *propensity score* dari kelompok perlakuan dan kontrol memiliki distribusi yang sama pada setiap kovariat.
6. Pengujian signifikansi dan interpretasi variabel *confounding*.

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.

Struktur Data Penelitian

Unit (i)	y_i	$x_{1.i}$	$x_{2.i}$	…	$x_{13.i}$
1	y_1	$x_{1.1}$	$x_{2.1}$	…	$x_{13.1}$
2	y_2	$x_{1.2}$	$x_{2.2}$	…	$x_{13.2}$
…	…	…	…	…	…
218	y_{218}	$x_{1.218}$	$x_{2.218}$	…	$x_{13.218}$

Variabel Penelitian

Variabel	Kategori
y: Status HIV/AIDS	0 =Negatif 1 =Positif
Variabel Identitas	
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

Variabel	Kategori
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

Variabel	Kategori
Riwayat Penggunaan Jarum Suntik	
x_9: Zat Yang Disuntikkan	1 = Putau 2 = Buphre 3 = Anti Depresan 4 = Campuran
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

$e(x_i)$

Estimasi Propensity Score Menggunakan Regresi Logistik

1. Fungsi distribusi probabilitas untuk setiap pasangan (x_i, z_i) yaitu

$$f(x_i) = e(x_i)^{z_i}\left(1 - e(x_i)\right)^{n-z_i} \quad (1)$$

$$e(x_i) = \frac{\exp\left[\left(\sum_{j=0}^{k}\beta_j x_{ij}\right)\right]}{1+\exp\left[\left(\sum_{j=0}^{k}\beta_j x_{ij}\right)\right]}, i = 1, 2, \cdots, n \quad \text{atau} \quad e(\mathbf{x}_i) = \frac{\exp\left[(\boldsymbol{\beta}'\mathbf{x}_i)\right]}{1+\exp\left[(\boldsymbol{\beta}'\mathbf{x}_i)\right]}$$

2. Fungsi likelihood

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

$$l(\boldsymbol{\beta}) = \left\{\exp\left[\sum_{i=1}^{n} z_i \log\left(\frac{e(\mathbf{x}_i)}{1-e(\mathbf{x}_i)}\right)\right]\right\}\left\{\prod_{i=1}^{n}\left(1-e(\mathbf{x}_i)\right)^{n}\right\} = \exp\left(\sum_{i=1}^{n} z_i\boldsymbol{\beta}'\mathbf{x}_i\right) - \sum_{i=1}^{n} n\log\left[1+\exp(\boldsymbol{\beta}'\mathbf{x}_i)\right] \quad (2)$$

$e(x_i)$

Estimasi Propensity Score Menggunakan Regresi Logistik

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

a Untuk kelompok perlakuan z_i=1

Berdasarkan Pers.(2)

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

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

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

$$\sum_{i=1}^{n} \mathbf{x}_{ia} - \sum_{i=1}^{n} n\hat{e}(\mathbf{x}_i)\mathbf{x}_{ia} = 0, \quad a = j = 0, 1, 2, \cdots, k; \quad \text{dengan} \quad \hat{e}(\mathbf{x}_i) = \frac{\exp\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_i\right)}{1 + \exp\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_i\right)}$$

$e(x_i)$

Estimasi Propensity Score Menggunakan Regresi Logistik

Turunan kedua dari fungsi log *likelihood*

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a^2} = -\sum_{i=1}^{n} n\mathbf{x}_{ia}^2 \hat{e}(\mathbf{x}_i)\left(1-\hat{e}(\mathbf{x}_i)\right)$$

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

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b} = -\sum_{i=1}^{n} \mathbf{x}_{ia}\mathbf{x}_{ib} n\hat{e}(\mathbf{x}_i)\left(1-\hat{e}(\mathbf{x}_i)\right); a,b = 0,1,2,\cdots,k.$$

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

$$Cov\left(\hat{\boldsymbol{\beta}}\right) = \left\{\mathbf{x}^T Diag\left[\hat{e}(\mathbf{x}_i)\left(1-\hat{e}(\mathbf{x}_i)\right)\right]\mathbf{x}\right\}^{-1}, \text{ dengan } \mathbf{x} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$e(x_i)$ Estimasi Propensity Score Menggunakan Regresi Logistik

Iterasi Newton Raphson

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - \left(\mathbf{H}^{(t)}\right)^{-1} \mathbf{q}^{(t)};$$

$$t = 1, 2, \cdots \textit{ sampai konvergen}$$

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

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

Untuk setiap langkah iterasi ke-t, berlaku:

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

$$q_j^{(t)} = \left. \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_j} \right|_{\beta^{(t)}} = \sum_{i=1}^{n} \left(1 - n e(\mathbf{x}_i)^{(t)}\right) \mathbf{x}_i$$

$$e(\mathbf{x}_i)^{(t)} = \frac{\exp\left(\boldsymbol{\beta}'^{(t)} \mathbf{x}_i\right)}{1 + \exp\left(\boldsymbol{\beta}'^{(t)} \mathbf{x}_i\right)}$$

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

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

$e(x_i)$ Estimasi Propensity Score Menggunakan Regresi Logistik

b Untuk kelompok kontrol z_i=0

Berdasarkan Pers.(2)

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

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

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

$$-\sum_{i=1}^{n} n\hat{e}\left(\mathbf{x}_i\right)\mathbf{x}_{ia} = 0,\ \ a = j = 0,1,2,\cdots,k;\ \ \text{dengan}\ \ \hat{e}\left(\mathbf{x}_i\right) = \frac{\exp\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_i\right)}{1+\exp\left(\hat{\boldsymbol{\beta}}'\mathbf{x}_i\right)}$$

$e(x_i)$ Estimasi Propensity Score Menggunakan Regresi Logistik

Turunan kedua dari fungsi log *likelihood*

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a^2} = -\sum_{i=1}^{n} n\mathbf{x}_{ia}^2 \hat{e}(\mathbf{x}_i)\left(1-\hat{e}(\mathbf{x}_i)\right)$$

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

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b} = -\sum_{i=1}^{n} \mathbf{x}_{ia}\mathbf{x}_{ib} n\hat{e}(\mathbf{x}_i)\left(1-\hat{e}(\mathbf{x}_i)\right); a,b = 0,1,2,\cdots,k.$$

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

$$Cov\left(\hat{\boldsymbol{\beta}}\right) = \left\{\mathbf{x}^T Diag\left[\hat{e}(\mathbf{x}_i)\left(1-\hat{e}(\mathbf{x}_i)\right)\right]\mathbf{x}\right\}^{-1} \quad \text{dengan} \quad \mathbf{x} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$e(x_i)$ Estimasi Propensity Score Menggunakan Regresi Logistik

Iterasi Newton Raphson

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - \left(\mathbf{H}^{(t)}\right)^{-1} \mathbf{q}^{(t)};$$

$$t = 1, 2, \cdots \textit{ sampai konvergen}$$

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

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

Untuk setiap langkah iterasi ke-t, berlaku:

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

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

$$e(\mathbf{x}_i)^{(t)} = \frac{\exp\left(\boldsymbol{\beta}'^{(t)} \mathbf{x}_i\right)}{1 + \exp\left(\boldsymbol{\beta}'^{(t)} \mathbf{x}_i\right)}$$

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

Dimana $m^{(t)} = ne(\mathbf{x}_i)^{(t)}$.

$e(x_i)$

Propensity Score pada Kasus HIV/AIDS

1 Statistik Deskriptif

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%)

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%.

Propensity Score Pada Kasus HIV/AIDS

$e(x_i)$

2 Pemilihan Variabel *Confounding*

H_0: Tidak terdapat hubungan signifikan.
H_1: Terdapat hubungan yang signifikan.

Taraf signifikansi: α=5%.

Statistik uji: $\chi^2 = \sum_{r=1}^{R}\sum_{c=1}^{C}\frac{(n_{rc}-e_{rc})^2}{e_{rc}}$.

Daerah kritik: H_0 ditolak jika

$\chi^2 > \chi^2_{1-\alpha}; df = (r-1)(c-1)$ atau P-*value*<α

Variabel	χ^2	Df	P-value	Keputusan
x_5*x_9	30,426	9	0,014	Tolak H_0
x_5*x_{10}	6,451	6	0,901	Gagal tolak H_0
x_5*x_{11}	4,236	3	0,307	Gagal tolak H_0
x_5*x_{12}	3,145	3	0,317	Gagal tolak H_0
x_5*x_{13}	6,340	3	0,343	Gagal tolak H_0

Variabel	χ^2	Df	P-value	Keputusan
x_6*x_9	7,313	3	0,031	Tolak H_0
x_6*x_{10}	6,170	4	0,254	Gagal tolak H_0
x_6*x_{11}	1,818	2	0,417	Gagal tolak H_0
x_6*x_{12}	0,251	1	0,707	Gagal tolak H_0
x_6*x_{13}	0,349	1	0,821	Gagal tolak H_0

Variabel	χ^2	Df	P-value	Keputusan
x_7*x_9	4,123	3	0,141	Gagal tolak H_0
x_7*x_{10}	4,239	6	0,644	Gagal tolak H_0
x_7*x_{11}	9,186	3	0,735	Gagal tolak H_0
x_7*x_{12}	7,422	1	0,027	Tolak H_0
x_7*x_{13}	0,001	1	0,053	Gagal tolak H_0

Variabel	χ^2	Df	P-value	Keputusan
x_8*x_9	4,249	3	0,236	Gagal tolak H_0
x_8*x_{10}	0,970	2	0,742	Gagal tolak H_0
x_8*x_{11}	4,007	1	0,083	Gagal tolak H_0
x_8*x_{12}	3,828	1	0,038	Tolak H_0
x_8*x_{13}	22,917	1	0,045	Tolak H_0

Confounding x_8 yaitu selalu pakai kondom

Selanjutnya, x_8 dinotasikan z dengan parameter τ

Propensity Score $e(x_i)$ pada Kasus HIV/AIDS

3 Model *Propensity Score*

$$e(x_i)=\frac{\exp\left(\begin{array}{l}-12,573+2,334x_{1}+0,047x_{2}-17,383x_{3.1}-2,063x_{3.2}+0,098x_{3.3}\\+4,274x_{3.4}-16,205x_{4.1}+2,961x_{4.2}+0,550x_{4.3}-1,122x_{5.1}-19,526x_{5.2}\\-0,734x_{5.3}+1,135x_{6.1}+1,159x_{6.2}+3,345x_{7.1}+1,880x_{7.2}-18,762x_{7.3}\\+1,846x_{9.1}-16,381x_{9.2}-15,921x_{9.3}-0,588x_{10.1}+0,301x_{10.2}\\+2,149x_{11.1}-0,690x_{11.2}+1,312x_{12}+1,127x_{13}\end{array}\right)}{1+\exp\left(\begin{array}{l}-12,573+2,334x_{1}+0,047x_{2}-17,383x_{3.1}-2,063x_{3.2}+0,098x_{3.3}\\+4,274x_{3.4}-16,205x_{4.1}+2,961x_{4.2}+0,550x_{4.3}-1,122x_{5.1}-19,526x_{5.2}\\-0,734x_{5.3}+1,135x_{6.1}+1,159x_{6.2}+3,345x_{7.1}+1,880x_{7.2}-18,762x_{7.3}\\+1,846x_{9.1}-16,381x_{9.2}-15,921x_{9.3}-0,588x_{10.1}+0,301x_{10.2}\\+2,149x_{11.1}-0,690x_{11.2}+1,312x_{12}+1,127x_{13}\end{array}\right)}$$

Propensity Score $e(x_i)$ pada Kasus HIV/AIDS

4

Evaluasi *Propensity Score*

Evaluasi *propensity score* dilakukan untuk mengetahui apakah kelompok perlakuan dan kontrol berditribusi sama atau *balance*.

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

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

Taraf signifikansi: α=5%

Statistik uji

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

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

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

$e(x_i)$

Propensity Score pada Kasus HIV/AIDS

Keputusan:

Variabel	KS_{hitung}	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
$x_{4(3)}$	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*).

Propensity Score pada Kasus HIV/AIDS

$e(x_i)$

5 Pengujian Signifikansi Variabel *Confounding* Selalu Pakai Kondom

Hipotesis

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

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

Taraf signifikansi: α=10%

Statistik uji:

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

Daerah kritis: H_0 ditolak jika p-value< α

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

Kesimpulan: Variabel *confounding* selalu pakai kondom signifikan.

Kesimpulan...

$e(x_i)$

1

Estimasi *propensity score* menggunakan regresi logistik dengan metode MLE, pada turunan pertama fungsi log *likehood* terhadap parameter β_a diperoleh:

Untuk kelompok perlakuan

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

Untuk kelompok kontrol

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

Karena hasil turunan pertama tidak *close form*, maka dilakukan iterasi Newton-Raphson

2

Penasun yang selalu menggunakan kondom dapat terkena HIV/AIDS 1,25 kali dibandingkan 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 penasun dapat berisiko tinggi terinfeksi HIV/AIDS berawal dari perlaku mereka dalam penggunaan jarum suntik lalu ketika penasun tersebut terinfeksi HIV/AIDS melalui jarum suntik maka hal tersebut dapat menularkannya kepada pasangan tetap ataupun pasangan tidak tetapnya melalui aktivitas seksual.

$e(x_i)$

Saran

- *Propensity score dengan model lain seperti regresi linier, CART, MARS dll*
- Variabel Confounding lebih dari satu
- Algorithm Pencocokan yang lain seperti NN, stratified, kernel dll

Penelitian selanjutnya...

DAFTAR PUSTAKA

Agresti, A. (1990). *Categorical Data Analysis*. New York: John Wiley and Sons, Inc.

Arniti, N. K. (2014). *Faktor-Faktor yang Berhubungan dengan Penerimaan Tes HIV oleh Ibu Hamil di Puskesmas Kota Denpasar*. Program Studi Ilmu Kesehatan Masyarakat, Universitas Udayana, Denpasar.

Becker, S. O., & Ichino, A. (2002). Estimation of Average Treatment Effect Based on Propensity Score. *The Stata Journal*, *2*(4), 358–377.

Besral, Utomo, B., & Zani, A. P. (2004). Potensi Penyebaran HIV dari Pengguna NAPZA Suntik ke Masyarakat Umum, *8*(2), 53–58.

D'Agostino, R. B. (1998). Tutorial In Biostatistics Propensity Score Method For Bias Reduction In The Comparison Of A Treatment To A Non-Randomized Control Group, *17*, 2265–2281.

Daniel, W. W. (1978). *Statistik Nonparametrik Terapan*. Jakarta: Gramedia.

Emsley, R., Lunt, M., Pickles, A., & Dunn, G. (2008). The Stata Journal. *The Stata Journal*, *8*(3), 334–353.

Freedman, D. A., & Berk, R. A. (2008). Weighting Regressions by Propensity Scores, (32), 392–409.

Gebel, M., & Voßemer, J. (2014). The impact of employment transitions on health in Germany. A difference-in-differences propensity score matching approach. *Social Science & Medicine (1982)*, *108*, 128–36.

Hosmer, D., & Lemeshow, S. (1989). *Applied Logistic Regression*. USA: John Wiley and Sons, Inc.

Imbens, G. W. (2007). The Role of the Propensity Score in Estimating Dose-Response Functions, *87*(3), 706–710.

Kambu, Y. (2012). *Analisis Faktor-Faktor Yang Mempengaruhi Tindakan Pencegahan Penularan HIV Oleh ODHA Di Sorong*. Tesis, Ilmu Keperawatan Kekhususan Keperawatan Medikal Bedah, Universitas Indonesia, Depok.

Komisi Penanggulangan AIDS Nasional. (2002). *Ancaman HIV/AIDS di Indonesia Semakin Nyata, Perlu Penanggulangan Lebih Nyata*. Jakarta.

Kristianti, S. (2012). Dukungan WPS dan Teman Pelanggan terhadap Penggunaan Kondom pada Pelanggan WPS di Semampir Kediri Shinta Kristiantii. *STIKES*, *5*(2).

Kumalasari, I. Y. (2013). *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.

Kurth, T., Walker, A. M., Glynn, R. J., Chan, K. A., Gaziano, J. M., Berger, K., & Robins, J. M. (2005). Results Of Multivariable Logistic Regression, Propensity Matching, Propensity Adjustment, and Propensity-Based Weighting Under Conditions Nonuniform Effect. *American Journal of Epidemiology*, *163*(3), 262–70.

Le, C. T. (1998). *Applied Categorical Data Analysis*. USA: John Wiley and Sons, Inc.

DAFTAR PUSTAKA

Li, H., Graham, D. J., & Majumdar, A. (2013). The Impacts of Speed Cameras on Road Accidents: An Application of Propensity Score Matching Methods. *Accident Analysis and Prevention*, *60*, 148–57.

Littnerova, S., Jarkovsky, J., Parenica, J., Pavlik, T., Spinar, J., & Dusek, L. (2013). Why to Use Propensity Score in Observational Studies? Case Study Based on Data from the Czech Clinical Database AHEAD 2006–09. *Cor et Vasa*, *55*(4), 383–390.

Lunceford, J. K., & Davidian, M. (2000). Stratification and Weighting Via the Propensity Score in Estimation of Causal Treatment Effects : A Comparative Study.

McCaffrey, D. F., Ridgeway, G., & Moral, A. R. (2004). Propensity Score Estimaton With Boosted Regression for Evaluating Causal Effect in Observational Studies. *Pschological Method*, *9*(4), 403.

Newgard, C. D., Hedges, J. R., Arthur, M., & Mullins, R. J. (2004). Advanced statistics: the propensity score--a method for estimating treatment effect in observational research. *Academic Emergency Medicine : Official Journal of the Society for Academic Emergency Medicine*, *11*(9), 953–61.

Pach, A., Wayne, W., & Praptoraharjo, I. (2006). *Laporan Penelitian Penyebaran HIV di Indonesia: Studi Etnografi tentang Jaringan Seksual dan Perilaku Berisiko Pemakai Napza Suntik (Penasun)*. Jakarta Pusat.

Praptoraharjo, I., Wiebel, W. W., Kamil, O., & Iii, A. P. (2007). Jaringan Seksual dan Perilaku Berisiko Pengguna Napza Suntik: Episode Lain Penyabaran HIV di Indonesia. *Berita Kedokteran Masyarakat*, *23*(3), 106–118.

Pratiwi, N. L., & Basuki, H. (2011). HIV-AIDS dan Perilaku Seks Tidak Aman di Indonesia. *Buletin Penelitian Sistem Kesehatan*, *14*(4), 346–357.

Ridgeway, G., McCaffrey, D., & Morral, A. (2011). Toolkit for Weighting and Analysis of Nonequivalent Groups : A Tutorial for the Twang Package.

Rosenbaum, P., & Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*(1), 41–55.

Rosenbaum, P., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, *79*(387), 516–524.

Suyasa, P. T. Y. S., & Wijaya, F. (2006). Resiliensi dan Sikap Terhadap Penyalahgunaan Zat (Studi Pada Remaja). *Jurnal Psikologi*, *4*(2).

Winarno, H., Suryoputro, A., & Shaluhiyah, Z. (2008). Faktor-Faktor Yang Berhubungan Dengan Penggunaan Jarum Suntik Bergantian Diantara Pengguna Napza Suntik Di Kota Semarang. *Jurnal Promosi Kesehatan Indonesia*, *3*(2).

Yanovitzky, I., Zanutto, E., & Hornik, R. (2005). Estimating Causal Effects of Public Health Education Campaigns using Propensity Score Methodology. *Evaluation and Program Planning*, *28*(2), 209–220.

$e(x_i)$

Terima Kasih