



PROPENSITY SCORE

Untuk Data Kesehatan

**Bambang Widjanarko Otok
Purhadi
Mahdalena**

PROPENSITY SCORE

Untuk Data Kesehatan

Penulis:

DR. Bambang Widjanarko Otok, Ir., M.Si
Dr. Purhadi, M.Sc
Dr. Mahdalena, S.Pd., M.Kes
ISBN.

Editor:

Eka Diah Kartiningrum, MKes

Penyunting:

Dr. Rifaatul Laila Mahmudah, MFarm.Klin.Apt

Desain Sampul dan Tata Letak:

Widya Puspitasari, AMd

Penerbit:

STIKes Majapahit Mojokerto

Redaksi:

Jalan Raya Jabon Km 02 Mojoanyar Mojokerto
Telp. 0321 329915
Fax. 0321 329915
Email: mojokertostikesmajapahit@gmail.com

Distributor Tunggal:

STIKes Majapahit Mojokerto
Jalan Raya Jabon Km 02 Mojoanyar Mojokerto
Telp. 0321 329915
Fax. 0321 329915
Email: mojokertostikesmajapahit@gmail.com

Cetakan pertama, November 2019

Hak Cipta Dilindungi undang-undang
Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan
cara apapun tanpa ijin tertulis dari penerbit

KATA PENGANTAR

Puji Syukur kehadiran Alloh SWT atas karunia dan hidayah Nya sehingga buku Propensity Score untuk data kesehatan dapat tersusun dengan baik. Buku ini merupakan dedikasi penulis terhadap kesejahteraan masyarakat di bidang kesehatan sehingga tujuan mulia pembangunan kesehatan di Indonesia dapat tercapai.

Rasa terima kasih tak lupa kami ucapkan untuk seluruh civitas Institut Teknologi Sepuluh November Surabaya yang telah memberikan dukungan moril maupun materiil sehingga buku yang merupakan intisari hasil penelitian yang penulis laksanakan dalam dua tahun terakhir dapat terrealisasi dengan baik.

Semua karya pasti ada kurang dan lebihnya, untuk kesempurnaan karya penulis selanjutnya, maka kritikan dan saran dari pembaca sangat kami harapkan. Semoga buku ini dapat bermanfaat bagi seluruh pembaca.

Surabaya, 19 November 2019

Tim Penulis

DAFTAR ISI

COVER	i
HALAMAN BALIK	ii
KATA PENGANTAR	iii
DAFTAR ISI	iv
CHAPTER 1 Pendahuluan	1
CHAPTER 2 Propensity Score Analysis	5
A. Konsep <i>Propensity Score</i>	5
B. Metode Berbasis <i>Propensity Score</i>	12
CHAPTER 3 Propensity Score Stratification <i>Using R Language</i>	33
CHAPTER 4 Propensity Score Weighting <i>Using R Language</i>	51
CHAPTER 5 Apply of Propensity Score with Marginal Mean Weighting Through Stratification <i>(MMWS) Using R Language</i>	65
CHAPTER 6 Propensity Score Stratification Bootstrap <i>Using R Language</i>	81
REFERENCES	105
LAMPIRAN	107
Riwayat Penulis	121

Pendahuluan

Metode penelitian yang biasa digunakan pada dasarnya disesuaikan dengan bidang yang akan diteliti. Pada bidang keilmuan eksakta atau sains seringkali digunakan metode penelitian yang bersifat eksperimen untuk mengetahui seberapa efektif suatu treatment yang diberikan. Dalam penelitian eksperimen ini subjek penelitian ditempatkan pada kelompok kontrol dan kelompok treatment berdasarkan pengacakan. Pada bidang sosial atau kesehatan penelitian eksperimen tidak selalu dapat digunakan, pada kondisi tertentu biasanya digunakan metode penelitian yang bersifat observasional atau non-eksperimental seperti survey, sensus, rekam medis pasien rumah sakit, dan lain-lain.

Penelitian non-eksperimental atau studi observasi merupakan investigasi empiris untuk mengetahui dampak dari suatu perlakuan pada saat eksperimen acak tidak layak dilakukan (Guo & Fraser, 2010). Jika dalam suatu penelitian pengacakan tidak dilakukan, maka akan muncul permasalahan bias yang disebabkan karena pemilihan subjek penelitian yang tidak acak. Permasalahan ini merupakan masalah seleksi bias yang dapat menyebabkan kesimpulan penelitian yang diperoleh tidak valid. Selain itu, untuk menganalisis hubungan sebab-akibat pada penelitian non-

eksperimental menggunakan model regresi *Ordinary Least Square* (OLS) dengan indikator perlakuan (dikotomus/politomus) tidak akan bekerja karena error dari model yang ada berkorelasi dengan variabel prediktor, hal ini tentunya dapat menyebabkan asumsi OLS tidak terpenuhi sehingga hasil estimasi efek perlakuan juga akan bias dan diragukan kebenarannya.

Oleh karena itu, dibutuhkan metode yang mampu mengatasi permasalahan seleksi bias diatas serta mampu menggambarkan hubungan sebab-akibat antara variabel prediktor dengan variabel respon pada penelitian non-eksperimental atau studi observasi. Salah satu metode yang dapat mengatasi permasalahan diatas adalah metode *propensity score*.

Metode *Propensity score* didefinisikan sebagai probabilitas bersyarat untuk menerima suatu treatment berdasarkan pada karakteristik-karakteristik sebelum treatment (Rosenbaum and Rubin, 1983). *Propensity score* diperkenalkan oleh Rosenbaum dan Rubin pada tahun 1983. Metode ini merupakan metode penyesuaian statistik yang dapat digunakan untuk menyeimbangkan atau menyamakan kelompok subjek penelitian sehingga dapat mengurangi bias akibat pemberian perlakuan yang tidak acak. Bias karena pengacakan ini akan tereduksi ketika hasil perbandingan antara kelompok control dan perlakuan hampir sama atau sedekat mungkin. Kata “terdekat” diukur dalam ukuran karakteristik-karakteristik yang dapat diobservasi (Akolo, 2016).

Seiring perkembangan zaman metode propensity score sudah banyak digunakan oleh para peneliti. Berbagai macam software yang digunakan untuk mengolah data berbasis propensity, misalnya software R dengan menggunakan package R yang sudah tersedia ataupun dengan menggunakan syntax manual, software STATA, software

SPSS dan software SAS. Pada buku ini akan dibahas secara rinci mengenai step by step tentang metode propensity score baik propensity score matching, propensity score stratification dan propensity score weighting menggunakan software R pada beberapa kasus dalam bidang kesehatan.

Halaman ini sengaja dikosongkan

Propensity Score Analysis

A. Konsep *Propensity Score*

Propensity Score pertama kali diperkenalkan oleh Rosenbaum dan Rubin tahun 1983 dalam jurnalnya yang berjudul “*The central role of the propensity score in observational studies for causal effects*”. Analisis *propensity score* adalah metode statistika inovatif yang berkembang pesat dan digunakan untuk mengevaluasi efek perlakuan pada data observasi (Guo & Fraser, 2010). Rosenbaum dan Rubin (1983) mendefinisikan *propensity score* untuk observasi ke- i ($i = 1, \dots, n$) sebagai probabilitas bersyarat dari *treatment* tertentu ($Z_i = 1$) berdasarkan pada karakteristik kovariat x_i yang diobservasi dimana pengacakan tidak bisa dilakukan. Nilai *propensity score* $e(\mathbf{x}_i)$ didefinisikan sebagai berikut (Guo & Fraser, 2010).

$$e(\mathbf{x}_i) = P(Z_i = 1 | X_i = x_i) \quad (1.1)$$

Propensity score merupakan metode yang tepat untuk mengatasi permasalahan seleksi bias pada data observasi. Metode *propensity score* pada umumnya ada tiga yakni: *Propensity Score Matching* (PSM), *Propensity Score Stratification* (PSS), dan *Propensity Score Weighting* (PSW). Rossi dan Freeman (1989) menjelaskan bahwa Metode PSM menitikberatkan bagaimana menyeimbangkan data kelompok

perlakuan dan kontrol dengan melakukan pencocokan berdasarkan kovariat yang diamati, sedangkan PSS merupakan prosedur dimana observasi dibagi menjadi beberapa kelompok atas dasar kovariat, dan PSW menitikberatkan pada pemberian bobot untuk setiap observasi berdasarkan kovariat yang diamati atau berdasarkan estimasi nilai dari *propensity score* (Guo & Fraser, 2010). Jika diterapkan dengan tepat, maka metode ini dapat membantu memecahkan masalah seleksi bias dan memberikan estimasi yang valid dari efek rata-rata perlakuan (Guo & Fraser, 2010).

Selain memecahkan masalah seleksi bias, salah satu keuntungan dari *propensity score matching* dan *stratification* adalah mereduksi dimensi, dimana jika vector \mathbf{x} memiliki banyak kovariat yang direpresentasikan dalam banyak dimensi maka pendekatan *propensity* mampu mereduksi semua dimensi menjadi satu skor dimensi, yakni nilai *propensity score* (Guo & Fraser, 2010).

Ada beberapa tahapan yang dilakukan dalam analisis propensity score (Yanovitzky, Zanutto, & Hornik, 2005), yakni:

1. Pemilihan kovariat sebagai variabel *confounding*

Confounding merupakan *distorsi* (gangguan) dalam menaksir pengaruh paparan terhadap kejadian penyakit/*outcome* akibat tercampurnya pengaruh faktor luar. Faktor luar (*extraneous variable*) ini yang dapat mempengaruhi faktor resiko dan penyakit secara bersama-sama dalam proses timbulnya penyakit, sehingga menyebabkan bias terhadap kesimpulan hasil studi. Faktor inilah yang disebut faktor *confounding*. Faktor *confounding* biasanya ada jika perbandingan antara kelompok perlakuan dan kontrol tidak seimbang. Kriteria variabel *confounding* yaitu: (1) Merupakan faktor resiko bagi kasus yang diteliti; (2)

Mempunyai hubungan dengan variabel bebas lainnya (Bruce, Pope, & Stanistreet, 2008). Proses pemilihan variabel *confounding* didasarkan pada teori dan bukti empiris tentang hubungan antar variabel. Adapun langkah-langkah untuk uji *confounding* adalah sebagai berikut

1. Melakukan pemodelan terhadap variabel penelitian
2. Melakukan uji interaksi antara variabel bebas dengan *confounding*
3. Melakukan uji signifikansi *confounder*
4. Interpretasi

Ada beberapa cara untuk mengontrol *confounding*, yaitu dengan mengontrol pada tahap design atau pada tahap analisis. Pada tahap design dilakukan dengan cara melakukan restriksi berupa membatasi obyek penelitian, pencocokan dan randomisasi sedangkan pada tahap analisis dilakukan dengan cara stratifikasi dan analisis multivariat (Bruce, et.al., 2008). Propensity score merupakan cara mengontrol *confounding* pada tahapan analisis yakni analisis multivariat.

Pemilihan *confounding* ini merupakan salah satu hal yang cukup penting dalam propensity score. Hal ini disebabkan karena model yang digunakan untuk estimasi nilai propensity pada tahapan selanjutnya bergantung pada tipe data dari kovariat dan *confounding* yang terpilih.

2. Mengidentifikasi Kovariat yang Tidak Balance Sebelum Propensity Score

Tahapan setelah pemilihan kovariat sebagai *confounder* adalah mengidentifikasi kovariat yang mempunyai potensi tidak

balance pada data penelitian. Kovariat yang tidak balance ini mengindikasikan bahwa terdapat bias pada data yang dapat menyebabkan kesimpulan tidak akurat. Proses identifikasi kovariat yang tidak balance dapat dilakukan menggunakan uji t dua sampel untuk treatment dengan dua level atau dengan uji One-Way Analysis Of Variance (ANOVA) untuk treatment yang lebih dari dua level. Kovariat yang tidak balance ini akan digunakan untuk membandingkan kovariat sebelum dilakukan propensity score dan setelah dilakukan propensity score.

Tahapan pemilihan variabel *confounding* dan identifikasi kovariat yang tidak balance dapat menjadi alasan yang kuat untuk memulai analisis propensity score. Akan tetapi, apabila dalam penelitian tidak ditemukan variabel yang berpotensi sebagai *confounding* dan tidak terdapat kovariat yang tidak balance, maka *propensity score* tidak perlu dilakukan.

3. Estimasi nilai propensity score

Tahapan ini salah satu tahapan yang sangat penting karena nilai yang dihasilkan merupakan nilai yang menjadi dasar pada proses selanjutnya. Metode yang digunakan untuk estimasi nilai propensity score bergantung pada tipe data dari kovariat dan variabel yang menjadi treatment dan kontrol. Nilai *propensity score* merupakan nilai yang menyatakan probabilitas dari subjek untuk terkena perlakuan berdasarkan kovariat \mathbf{x}_i dan biasa disimbolkan dengan $e(\mathbf{x}_i)$.

Beberapa penelitian tentang model yang digunakan untuk estimasi nilai *propensity score*, yakni: McCandless, Gustafson & Austin (2009) menggunakan metode *bayesian* untuk estimasi nilai propensity, McCaffrey, Ridgeway & Moral (2004) yang menggunakan *generalized boosted model*, dan Littnerova, Jarkovsky, Parenica, Pavlik, Spinar & Dusek (2013)

menggunakan regresi logistik untuk estimasi *propensity score*. Ketiga penelitian tersebut diterapkan pada data yang bersifat kategori dengan treatment terdiri atas dua kategori. Untuk penelitian yang menggunakan variabel laten maka model *Structural Equation Modelling* (SEM) merupakan salah satu model untuk estimasi nilai propensity score yang tepat. Akan tetapi, sebagian besar penelitian yang ada biasanya menggunakan model regresi logistik untuk estimasi nilai propensity. Hal ini disebabkan karena model regresi logistik tidak membutuhkan metode yang *complicated* untuk memperoleh estimasi parameternya khususnya parameter nilai propensity score.

Menurut Littnerova, et.al., (2010), nilai propensity score untuk treatment yang terdiri dua level dinyatakan dalam model regresi logistik sebagai berikut.

$$e(\mathbf{x}_i) = P(Z_i = 1 | X_i = x_i) = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})} \quad (1.2)$$

dengan β_0 merupakan konstanta, $\beta_1, \beta_2, \dots, \beta_p$ merupakan koefisien regresi dan x_1, x_2, \dots, x_p merupakan variabel kovariat.

4. Menggunakan Metode Propensity Score

Setelah diperoleh nilai propensity score, maka langkah selanjutnya adalah menggunakan metode propensity score yang ada untuk diterapkan pada kelompok treatment dan kelompok kontrol. Tahapan ini bergantung pada metode propensity apa yang digunakan, misalnya jika menggunakan propensity score matching maka dilakukan pencocokkan

antara kelompok treatment dan kontrol, jika digunakan propensity score stratification maka dilakukan pembagian subjek ke dalam strata, atau jika digunakan propensity score weighting maka dilakukan pemberian bobot pada subjek kelompok treatment dan kontrol. Tahapan inilah yang dapat menyeimbangkan kelompok treatment dan kontrol sehingga bias karena *confounding* atau masalah seleksi bias dapat direduksi.

5. Memeriksa Balance Setiap Kovariat.

Seperti yang telah dijelaskan sebelumnya bahwa propensity score merupakan metode yang mampu mereduksi bias pada data observasi karena faktor *confounding*. Bias pada data observasi dapat direduksi dengan cara menyeimbangkan distribusi kelompok treatment dan kontrol. ketika kelompok treatment dan kelompok kontrol memiliki distribusi yang sama, misalnya mean/ rata-rata, proporsi antar kelompok treatment dan kontrol sama maka berarti bias yang ada telah tereduksi. Keseimbangan dalam propensity score dikenal dengan istilah *balance*. *Balance* memberikan gambaran bahwa tidak ada perbedaan distribusi antara kelompok treatment dan kontrol atau tidak ada perbedaan yang signifikan antara kelompok treatment dan kontrol.

Pada beberapa penelitian sebelumnya, ada beberapa statistik uji yang digunakan untuk memeriksa balance antara kelompok treatment dan kelompok kontrol. Statistik uji yang digunakan bergantung pada tipe data yang digunakan. Untuk tipe data kontinu biasanya menggunakan statistic uji t untuk menguji perbedaan mean/rata-rata antara kelompok treatment dan kontrol atau dapat juga menggunakan uji Kolmogorov-Smirnov untuk membuktikan bahwa kelompok treatment dan kontrol memiliki distribusi yang sama. Untuk tipe data kategori

biasanya menggunakan statistic uji chi-square atau uji-z untuk menguji perbedaan proporsi antara kelompok treatment dan kontrol. Pengecekan balance ini dilakukan untuk setiap kovariat yang ada. Apabila terdapat kovariat yang tidak *balance* atau dalam artian terdapat perbedaan antara kelompok treatment dan kontrol maka hal ini mengindikasikan bahwa tahapan tidak bisa dilanjutkan karena asumsi dasar dari metode propensity score tidak terpenuhi. Metode ini dapat dilanjutkan jika asumsi balance untuk semua kovariat terpenuhi.

6. Menghitung Efek Treatment

Estimasi *Average Treatment Effect* (ATE) diformulasikan dengan rumus berikut

$$\hat{\tau} = E(\hat{y}_1 | Z_i = 1) - E(\hat{y}_0 | Z_i = 0) \quad (1.3)$$

dengan $\hat{\tau}$ merupakan Estimasi *Average Treatment Effect* (ATE), \hat{y}_1 merupakan respon untuk kelompok treatment, \hat{y}_0 merupakan respon untuk kelompok kontrol, $Z_i = 1$ merupakan kelompok treatment, $Z_i = 0$ merupakan kelompok control.

Estimasi *Average Treatment Effect on Treated* (ATT) diformulasikan dengan rumus berikut

$$\hat{\tau} = E(\hat{y}_1 - \hat{y}_0 | Z_i = 1) \quad (1.4)$$

dengan $\hat{\tau}$ merupakan Estimasi *Average Treatment Effect* (ATE), \hat{y}_1 merupakan respon untuk kelompok treatment, \hat{y}_0 merupakan respon untuk kelompok kontrol, $Z_i = 1$ merupakan kelompok treatment, $Z_i = 0$ merupakan kelompok control.

7. Evaluasi Keباikan Propensity Score

Untuk mengetahui kebaikan dari metode propensity score, perlu dilakukan pengujian untuk mengetahui apakah metode propensity score yang digunakan sudah mampu memberikan hasil yang diinginkan atau tidak. Kriteria kebaikan metode Propensity score dapat dilihat dari seberapa besar bias yang mampu direduksi oleh metode tersebut untuk setiap kovariat maupun bias secara keseluruhan. Untuk mengetahui besar bias yang direduksi, ada beberapa formulasi yang digunakan tergantung metode propensity yang digunakan oleh peneliti. Formulasi bias ini dibahas pada sub-bab yang membahas masing-masing metode propensity score.

B. Metode Berbasis Propensity Score

Menurut Guo & Fraser (2010) secara umum ada 3 metode yang berbasis propensity score, yakni metode *Propensity Score Matching* (PSM), *Propensity Score Stratification* (PSS) dan *Propensity Score Weigthing* (PSW). Namun seiring perkembangan zaman, telah dikembangkan metode baru yang digunakan untuk mengatasi kelemahan dari metode propensity sebelumnya. Metode ini adalah metode *Propensity Score Stratification Bootstrap* (PSSB), metode propensity score dengan *Marginal Mean Weighting Through Stratification* (MMWS). Pada sub-bab ini akan dibahas lebih detail tentang metode-metode yang berbasis propensity score.

1. *Propensity Score Matching* (PSM)

Salah satu metode propensity score yang banyak digunakan adalah propensity score matching. Propensity score matching adalah metode statistik yang melakukan pencocokan antara satu subjek treatment dengan satu atau

lebih subjek kontrol berdasarkan karakteristik kelompok yang diobservasi. Setiap subjek kelompok yang terkena treatment dipasangkan dengan subjek kelompok kontrol yang paling sama dengan probabilitas kelompok treatment atau nilai propensity score. Pencocokan ini membantu memperkuat pernyataan hubungan dalam studi non-eksperimental atau studi observasi. Propensity score matching paling bagus digunakan jika banyaknya subjek pada kelompok kontrol lebih banyak dibandingkan subjek pada kelompok treatment. Oleh karena propensity matching menitikberatkan pada pencocokan antara treatment dan kontrol, maka subjek yang tidak cocok akan dieliminasi. Sehingga yang akan digunakan pada analisis selanjutnya hanya subjek yang cocok saja (match).

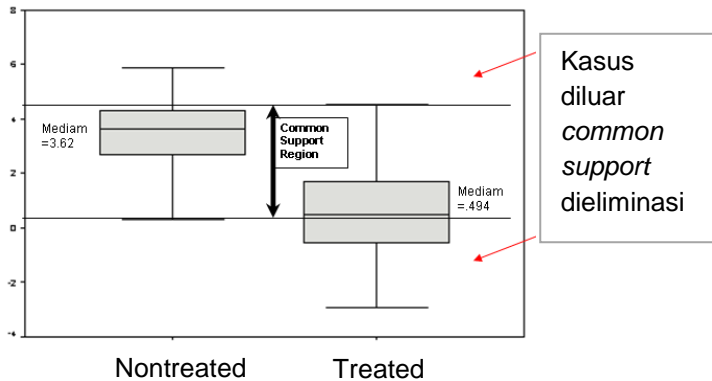
Propensity score matching terdiri atas dua kelompok besar, yakni *greedy matching* dan *optimal matching*.

a. Greedy Matching

Ide awal dari propensity score matching adalah membentuk sampel baru yang mempunyai kesamaan sesuai dengan kondisi treatment. Tahapan seperti ini umumnya disebut *Greedy Matching*. Kelebihan dari metode propensity score menggunakan *Greedy Matching* adalah (1) adanya analisis post-matching dan metode nearest neighbor within caliper matching yang sangat berguna jika variabel respon (outcome) adalah variabel yang tidak normal dan tidak kontinu misalnya variabel dependen dengan tipe kategorikal; (2) Mahalanobis metric matching dapat digunakan untuk mengetahui pasangan yang cocok ketika sampel yang digunakan kecil. oleh karena itu metode *Greedy Matching* ini sangat populer dibandingkan metode optimal matching.

Ada beberapa kelemahan dari metode *Greedy Matching*, yakni:

- *Greedy matching* mengasumsikan adanya overlapping estimasi propensity score antara kelompok treatment dan kelompok kontrol. Overlapping ini menunjukkan bahwa tidak terdapat perbedaan antara kelompok treatment dan kelompok kontrol. Oleh karena itu, subjek yang tidak cocok tidak diikutsertakan pada proses selanjutnya. Untuk lebih jelasnya dapat dilihat pada bagan berikut ini.



- *Greedy matching* dikritik karena mensyaratkan daerah *common support* yang lebih besar. Ketika daerah tersebut tidak ada pada kasus, maka *greedy matching* gagal diterapkan.

Greedy Matching terdiri atas beberapa metode yakni:

1. Nearest Neighbor Matching

Pada metode ini, absolut untuk perbedaan estimasi propensity score untuk kelompok treatment dan kontrol diminimumkan. Subjek treatment dan kontrol diurutkan secara acak, kemudian dipilih subjek treatment pertama yang paling

dekat nilai propensity score-nya dengan subjek control. Secara matematis prinsip *nearest neighbor* dituliskan sebagai berikut.

$$C(e(x_i)) = \min |e(x_i) - e(x_j)| \quad (1.5)$$

dengan $C(e(x_i))$ merupakan grup dari subjek kontrol ke- j yang dicocokkan dengan subjek treatment ke- i berdasarkan nilai propensity score, $e(x_i)$ merupakan estimasi propensity score untuk subjek treatment ke- i , dan $e(x_j)$ merupakan estimasi propensity score untuk subjek treatment ke- j .

2. *Radius Matching*

Pada metode radius matching, setiap subjek treatment diasumsikan cocok dengan subjek kontrol yang termasuk dalam interval dari nilai propensity score subjek treatment yang telah didefinisikan sebelumnya. Oleh karena subjek treatment harus cocok dengan subjek kontrol sesuai interval yang diberikan, maka hanya ada sejumlah tertentu dari perbandingan yang tersedia.

3. *Kernel Matching*

Pada metode ini, setiap subjek treatment dicocokkan dengan rata-rata bobot dari subjek kontrol. Bobot yang dimaksudkan berbanding terbalik dengan jarak antara nilai propensity score kelompok treatment dan kontrol.

4. *Caliper Matching*

Pada metode ini ditentukan terlebih dahulu range nilai yang biasanya didefinisikan dengan $\frac{1}{4}$ dari standar error (0,25

standar error) dari estimasi propensity. Semua nilai yang berada diluar range ini dikeluarkan dari analisis (Sianesi, 2002)

5. *Stratification Matching*

Metode *stratification matching* merupakan prosedur yang menggabungkan *matching* dan *stratification*. Langkah awal dari metode ini adalah melakukan pencocokan antara subjek treatment dan kontrol berdasarkan nilai propensity. Selanjutnya data yang cocok, dibagi ke dalam strata-strata berdasarkan urutan nilai propensity score. Setiap interval terdiri dari perlakuan dan kontrol subyek yang rata-rata memiliki skor kecenderungan setara.

Perbedaan antara hasil dari perlakuan dan kelompok kontrol dihitung untuk mendapatkan efek pengobatan rata-rata. Ini adalah rata-rata hasil dari pengobatan per blok ditimbang dengan distribusi subjek diperlakukan di blok.

6. *Mahalanobis Metric Matching*

Menurut Cochran & Rubin (1973) dalam Guo & Fraser (2010) mahalanobis metric matching merupakan metode yang sudah ada sebelum metode propensity score matching. Untuk mengaplikasikan metode ini, terlebih dahulu subjek diurutkan dan dihitung jarak antara subjek treatment pertama dengan semua subjek kontrol. Jarak Mahalanobis untuk jarak $d(i, j)$, antara subjek treatment ke- i dan subjek kontrol ke- j didefinisikan sebagai berikut.

$$d(i, j) = (\mathbf{u} - \mathbf{v})^T C^{-1} (\mathbf{u} - \mathbf{v}) \quad (1.6)$$

dengan \mathbf{u} dan \mathbf{v} adalah vektor nilai dari variabel yang matching untuk subjek treatment ke- i dan subjek kontrol ke- j , dan \mathbf{C} adalah matriks kovarian sampel dari variabel yang matching untuk semua subjek kontrol. subjek treatment dan kontrol dicocokkan berdasarkan nilai terkecil dari jarak Mahalanobis. Proses ini diulang sampai setiap subjek treatment sudah dicocokkan dengan kontrol, sedangkan untuk subjek kontrol yang tidak cocok dikeluarkan dari analisis.

Jika subjek treatment tidak memiliki nilai propensity score yang serupa dengan subjek kontrol, maka kesimpulan hubungan sebab akibat tidak dapat dilakukan tanpa menggunakan ekstrapolasi. Oleh karena itu, unit tersebut umumnya dikeluarkan dari analisis. Mahalanobis matching setelah propensity score matching diasumsikan setara dengan blocking dalam *randomized experimental design*.

b. Optimal Matching

Optimal Propensity Score Matching (OPSM) hanya dikenal selama 10 tahun saja. Meskipun begitu aplikasi dari OPSM telah berkembang pesat dan sering dimanfaatkan dalam penelitian. Dua alasan penggunaan OPSM, yakni: (1) penggunaan *network flow theory* untuk mengoptimalkan matching dan ketersediaan software komputasi yang cepat dalam hal ini package R **optmath** membuat OPSM layak digunakan. Rosenbaum (2002) menjelaskan secara komprehensif review dari teori dan prinsip aplikasi dari *optimal matching*.

Metode OPSM bergantung pada struktur data dari variabel penelitian. OPSM dapat diklasifikasikan menjadi 3 tipe, yakni:

- a. *Pair Matching*, dimana setiap subjek treatment dicocokkan dengan satu subjek kontrol.
- b. *Matching using a variabel ratio or variable matching*, dimana setiap subjek treatment dicocokkan dengan minimal satu sampai 4 subjek kontrol.
- c. *Full matching*, dimana setiap subjek treatment dicocokkan dengan satu atau lebih subjek kontrol dan sebaliknya setiap subjek kontrol dicocokkan dengan satu atau lebih subjek treatment.

OPSM merupakan proses dari mengembangkan himpunan yang dicocokkan ($A_1, \dots, A_s; B_1, \dots, B_s$) dengan ukuran (α, β) dimana total jarak dari nilai propensity score diminimumkan. Definisi minimum optimal matching dari total jarak adalah sebagai berikut.

$$\Delta = \sum_{k=1}^K \omega(|A_k|, |B_k|) \delta(A_k, B_k) \quad (1.7)$$

dengan $\omega(|A_k|, |B_k|)$ merupakan fungsi bobot, k merupakan strata dimana $k = 1, 2, \dots, K$, dan δ merupakan perbedaan antara treatment dan kontrol dari kovariat yang diobservasi, misalnya perbedaan berdasarkan nilai propensity score atau *Mahalanobis metrics*.

Tahapan dari Optimal Propensity Score Matching (OPSM) adalah sebagai berikut.

- a. Run R ***optmatch***: berdasarkan rasio dari sampel treatment. Memutuskan perbedaan struktur matching yang digunakan, apakah pair matching, full matching atau variable matching with different treated-to-control ratios.

- b. Cek kovariat imbalance menggunakan metode yang disarankan oleh Haviland et al. (2007) dan diharapkan setelah matching sampel treatment dan kontrol sudah balance untuk semua kovariat yang diobservasi.
- c. Post-matching analysis, berdasarkan sampel yang cocok, run analisis outcome untuk menghitung efek treatment.

2. Propensity Score Stratification (PSS)

Dalam studi observasi yang mempelajari hubungan sebab akibat, treatment merupakan unit eksperimen tanpa mempertimbangkan pengacakan. Hal ini menyebabkan kelompok treatment yang ada kemungkinan secara sistematis berbeda dengan karakteristik yang relevan sehingga tidak dapat dibandingkan dengan kelompok kontrol. Untuk mengontrol perbedaan sistematis ini, maka dilakukan pengelompokan subjek ke dalam kelas-kelas berbeda berdasarkan karakteristik yang diamati dan kemudian dibandingkan antara kelompok treatment dengan kelompok kontrol untuk setiap sub-kelas yang sama. Prosedur tersebut diyakini dapat mengontrol bias pada studi observasi dan dikenal dengan metode *Propensity Score Stratification* (PSS).

Propensity Score Stratification (PSS) merupakan salah satu metode propensity score yang diperkenalkan oleh Rosenbaum dan Rubin pada tahun 1984. Metode PSS adalah suatu prosedur untuk mengelompokkan subjek ke dalam strata atau kelas-kelas berdasarkan kovariat x_i yang diamati atau berdasarkan estimasi *propensity score* (Rosenbaum, 2002b). Subjek yang akan distratifikasi diurutkan menurut estimasi propensity score (Austin, 2011).

Untuk memenuhi asumsi balance kovariat pada propensity score, diperkenalkan prosedur stratifikasi berdasarkan persentil dari x_i yang disebut *quantile stratification*. Menurut Guo & Fraser (2010), *prosedur quantile stratification* didesain untuk variabel x dengan tipe data kontinu. Prosedur *Quantile stratification* mengikuti 5 tahapan berikut.

1. Mengurutkan data pada variabel x dari nilai terkecil sampai nilai terbesar berdasarkan urutan x
2. Jika akan dibuat 5 stratum maka ada 5 kuantil yang terbentuk. Selanjutnya memilih subjek yang nilai x sama atau kurang dari kuantil pertama untuk ditempatkan pada stratum pertama. Misalkan stratum 1 terdiri dari semua subjek yang memiliki nilai x sama dengan atau kurang dari nilai persentil ke-20 dari x . Selanjutnya memilih subjek yang nilai x berada pada rentang kuantil kedua, dalam hal ini stratum 2 terdiri dari semua subjek yang nilai x berada pada rentang nilai persentil ke-21 sampai persentil ke-40. Hal ini dilakukan sampai diperoleh 5 stratum.
3. Menghitung perbedaan mean respon antara kelompok treatment dan kontrol pada setiap stratum k ($k = 1, 2, \dots, 5$). Formulasi yang digunakan untuk menghitung perbedaan mean adalah sebagai berikut.

$$\delta_k = E(y_{k1} | Z_k = 1) - E(y_{k0} | Z_k = 0); \quad \text{dengan } k = 1, 2, \dots, 5 \quad (1.8)$$

dengan δ_k merupakan beda mean kelas ke- k , $E(y_{k1} | Z_k = 1)$ merupakan mean respon untuk treatment kelas ke- k dan $E(y_{k0} | Z_k = 0)$ merupakan mean respon untuk treatment kelas ke- k .

4. Menghitung mean dari rata-rata 5 strata untuk memperoleh *Average Treatment Effect (ATE)* menggunakan persamaan (1.6).

$$\hat{\tau} = \sum_{k=1}^K \frac{n_k}{N} (\bar{Y}_{k1} - \bar{Y}_{k0}) \quad (1.9)$$

dengan $\hat{\tau}$ merupakan estimasi ATE, n_k merupakan banyaknya subjek pada kelas ke K, N merupakan banyak subjek secara keseluruhan, \bar{Y}_{k1} dan \bar{Y}_{k0} merupakan proporsi respon untuk treatment dan kontrol kelas ke-k.

5. Menghitung varians dari estimasi *Average Treatment Effect* (ATE) menggunakan persamaan (1.7).

$$\text{Var}(\hat{\tau}) = \sum_{k=1}^K \left(\frac{n_k}{N} \right)^2 \text{Var}(\bar{Y}_{k1} - \bar{Y}_{k0}) \quad (1.10)$$

6. Uji signifikansi efek treatment

Pengujian efek treatment menggunakan uji z dua sampel karena treatment merupakan variabel yang bertipe kategori.

Ada beberapa hal yang menjadi permasalahan serius pada metode PSS, yakni pemilihan banyaknya strata yang optimal. Strata yang optimal berarti strata yang memenuhi asumsi balance untuk setiap kovariat yang ada serta strata yang mampu memberikan reduksi bias terbesar. Cochrane (1968) merepresentasikan satu contoh kasus kematian antara perokok dan bukan perokok. Dalam penelitiannya Cochrane membandingkan kelompok perokok dan bukan perokok untuk kovariat umur sebelum dan setelah dilakukan stratifikasi. Hasilnya menunjukkan bahwa 5 sub-kelas seringkali sudah cukup mampu mereduksi hamper 90% dari bias dengan satu kovariat tunggal (Mingxiang, 2012). Imbens (2004) menyatakan seluruh bias di bawah *unconfounded* diasosiasikan dengan *propensity score*, hal ini menunjukkan bahwa di bawah normalitas penggunaan 5 strata mengubah sebagian besar bias dengan seluruh kovariat.

Asumsi balance kovariat tidak selamanya dapat dipenuhi. Pada beberapa kondisi terkadang ditemui kasus yang *imbalance* dan membutuhkan solusi supaya tahapan dapat dilanjutkan sampai estimasi efek treatment. Menurut Dehejia & Wahba (1999) jika setelah dilakukan stratifikasi diperoleh kovariat yang tidak seimbang maka ada 2 cara yang dapat dilakukan, yakni: (1) kovariat yang tidak seimbang memberikan gambaran bahwa spesifikasi nilai propensity score yang diperoleh sebelumnya tidak memadai dan perlu ditentukan kembali, misalnya dengan menambahkan orde yang lebih tinggi dalam model berupa interaksi antara kovariat; (2) membagi lagi subjek ke dalam sub-kelas sampai diperoleh strata yang memenuhi asumsi balance kovariat.

Seperti yang telah dijelaskan sebelumnya bahwa salah satu kriteria kebaikan dari metode propensity score adalah dilihat dari bias yang direduksi oleh metode tersebut. Hal yang sama juga berlaku untuk metode propensity score stratification. Semakin besar bias yang mampu direduksi oleh metode propensity, maka semakin baik pula metode tersebut untuk menyelesaikan masalah bias pada kasus studi observasi.

Menurut Cochran & Rubin (1973) dalam Pan & Bai (2015) mengukur besar bias yang direduksi untuk masing-masing kovariat dapat menggunakan persamaan (1.11) sampai (1.14).

$$PBR = \frac{B_{\text{before PS}} - B_{\text{after PS}}}{B_{\text{before PS}}} \times 100\% \quad (1.11)$$

dan

$$B_{\text{kontinu}} = \bar{x}_{p1} - \bar{x}_{p0} \quad (1.12)$$

$$B_{\text{kategori}} = p_1(x_p) - p_0(x_p) \quad (1.13)$$

$$B_{\text{after PSS}} = \sum_{k=1}^K B_k \quad (1.14)$$

dengan PBR merupakan *Percent Bias Reduction*, B_{kontinu} merupakan selisih mean dari kelompok perlakuan dan

kelompok kontrol untuk setiap kovariat kontinu dan $B_{kategori}$ merupakan selisih proporsi untuk kovariat kategori, \bar{x}_{p1} dan \bar{x}_{p0} masing-masing merupakan mean kovariat untuk kelompok perlakuan dan kelompok kontrol, $p_1(x_p)$ dan $p_0(x_p)$ masing-masing merupakan proporsi kovariat untuk kelompok perlakuan dan kelompok kontrol, $B_{before PS}$ dan $B_{after PS}$ masing-masing merupakan selisih rata-rata kelompok perlakuan dan kontrol sebelum dilakukan *propensity score* dan setelah dilakukan *propensity score*.

3. *Propensity Score Weighting (PSW)*

Propensity score weighting merupakan salah satu metode propensity yang cukup populer dan lebih robust dibandingkan metode propensity lainnya. Metode ini menggunakan pembobot pada observasi ketika melakukan estimasi pengaruh perlakuan (McCaffrey, Ridgeway, & Moral, 2004). Propensity score weighting bertujuan melakukan pembobotan kembali pada subjek kelompok treatment dan kontrol untuk membuat populasi yang lebih representative. Pembobotan dilakukan berdasarkan estimasi nilai propensity score.

Menurut Guo & Fraser (2010) formulasi bobot yang digunakan pada propensity score weighting adalah sebagai berikut.

$$w_i = \frac{e(\mathbf{x}_i)}{(1 - e(\mathbf{x}_i))} \quad (1.15)$$

dengan w_i merupakan bobot untuk observasi ke- i dan $e(\mathbf{x}_i)$ merupakan nilai propensity score untuk observasi ke- i .

Untuk estimasi $E(y_0 | z=1)$, misalkan pengamatan i pada sampel pembandingan memiliki bobot w_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. Rata-rata bobot dari hasil pengamatan pada kelompok kontrol adalah

$$\hat{E}(y_0 | z=1) = \frac{\sum_{i \in C} w_i y_i}{\sum_{i \in C} w_i}, \quad (1.16)$$

dimana $i \in C$ menyatakan pengamatan ke- i pada kelompok kontrol dan penjumlahan pengamatan pada kelompok kontrol. $\hat{E}(y_0 | z=1)$ menyatakan estimasi nilai ekspektasi. Misalkan N_T menyatakan banyaknya pengamatan pada kelompok perlakuan dan $i \in T$ menyatakan pengamatan ke- i pada kelompok perlakuan, rata-rata sampel untuk pengamatan adalah $\hat{E}(y_1 | z=1) = \sum_{i \in T} y_i / N_T$.

Perbedaan tipe bobot yang digunakan pada metode ini bergantung pada efek treatment yang akan diestimasi, dalam hal ini *Average Treatment Effect* (ATE) dan *Average Treatment Effect on Treated* (ATT). Berikut ini ketentuan penggunaan bobot untuk estimasi ATE dan ATT (Guo & Fraser, 2010).

- Untuk estimasi ATE, bobot didefinisikan sebagai berikut.

$$\omega(W_i, x_i) = \frac{W_i}{e(x_i)} + \frac{1-W_i}{1-e(x_i)} \quad (1.17)$$

Berdasarkan definisi ini, ketika $W_i = 1$ (untuk kelompok treatment) maka persamaan (1.17) akan menjadi

$$\omega(W_i, x_i) = \frac{1}{e(x_i)}, \text{ dan ketika } W_i = 0 \text{ (untuk kelompok}$$

control) maka persamaan (1.17) akan menjadi

$$\omega(W_i, x_i) = \frac{e(\mathbf{x}_i)}{(1 - e(\mathbf{x}_i))}$$

- Untuk estimasi ATT, bobot didefinisikan sebagai berikut.

$$\omega(W_i, x_i) = W_i + (1 - W_i) \frac{e(x_i)}{1 - e(x_i)} \quad (1.18)$$

Berdasarkan definisi ini, ketika $W_i = 1$ (untuk kelompok treatment) maka persamaan (1.18) akan menjadi $\omega(W_i, x_i) = 1$, dan ketika $W_i = 0$ (untuk kelompok control)

maka persamaan (1.18) akan menjadi $\omega(W_i, x_i) = \frac{e(\mathbf{x}_i)}{(1 - e(\mathbf{x}_i))}$

4. *Propensity Score Stratification Bootstrap (PSSB)*

Bootstrap merupakan salah satu metode resampling yang digunakan untuk mengukur tingkat akurasi ukuran statistic misalnya bias dan standar error suatu parameter serta dapat menaksir parameter-parameter pada suatu distribusi. Metode bootstrap pertama kali diperkenalkan oleh Efron pada tahun

1979. Dalam perkembangannya metode ini sudah banyak dikombinasikan dengan metode lainnya, salah satunya dengan metode propensity score. Metode bootstrap diyakini dapat mengatasi permasalahan-permasalahan yang muncul dalam metode propensity score khususnya jika berkaitan dengan asumsi independensi strata yang tidak terpenuhi dan jumlah sampel yang kecil. Hal ini disebabkan karena bootstrap memiliki kelebihan yang dapat mengatasi permasalahan statistika tanpa perhitungan yang rumit, tanpa asumsi yang harus dipenuhi karena hanya didasarkan pada data yang ada. Ada dua metode propensity score yang dikombinasikan dengan bootstrap. Metode tersebut adalah *Propensity Score Matching Bootstrap* (PSMB) dan *Propensity Score Stratification Bootstrap* (PSSB).

Propensity Score Stratification Bootstrap (PSSB) pada dasarnya tidak jauh berbeda dengan *Propensity Score Stratification* tanpa bootstrap. Tahapan yang digunakan sama, akan tetapi tahapan untuk PSSB lebih panjang dibandingkan PSS tanpa bootstrap. Hal ini disebabkan karena untuk PSSB tahapan propensity dilakukan 2 kali, yakni untuk full data dan untuk sampel bootstrap. Pada tahap pengujian, ada sedikit perbedaan antara pengujian parameter untuk full data dan sampel bootstrap. Untuk menguji signifikansi parameter pada full data digunakan uji Z sesuai persamaan (1.19) sedangkan untuk menguji signifikansi parameter bootstrap digunakan uji Z sesuai persamaan (1.20)

$$Z_0 = \frac{\hat{\beta}_p}{SE(\hat{\beta}_p)} \quad (1.19)$$

dan

$$Z^* = \frac{\hat{\beta}_p^* - \hat{\beta}_p}{SE(\hat{\beta}_p^*)} \quad (1.20)$$

dengan Z_0 digunakan untuk menguji signifikansi parameter sebelum *bootstrap* dilakukan, Z^* digunakan untuk menguji signifikansi parameter setelah *bootstrap* dilakukan, $\hat{\beta}_p$ merupakan parameter pada data tanpa *bootstrap*, $\hat{\beta}_p^*$ merupakan parameter pada sampel *bootstrap*, $SE(\hat{\beta}_p)$ merupakan standar error dari parameter data tanpa *bootstrap* dan $SE(\hat{\beta}_p^*)$ merupakan standar error dari parameter data sampel *bootstrap* untuk variabel prediktor $p = 1, 2, \dots, P$ (Davison & Hinkley, 1997).

Daerah penolakan jika menggunakan persamaan (1.19) yakni tolak H_0 jika $|Z_0| > Z_{\alpha/2}$ atau $p\text{-value} < \alpha$. Daerah penolakan H_0 untuk parameter *bootstrap* berbeda dengan daerah penolakan pada parameter tanpa *bootstrap*. Daerah penolakannya menggunakan $p\text{-value}$ dimana tolak H_0 jika $p\text{-value} < \alpha$. $p\text{-value}$ yang digunakan ditunjukkan oleh persamaan (2.7).

$$p\text{-value} = \frac{1 + \#(Z^* \geq Z_0)}{R + 1} \quad (1.21)$$

dengan

$$\#(Z^* \geq Z_0) = \sum_{r=1}^R I(Z^* \geq Z_0) \quad (1.22)$$

Z_0 merupakan statistic uji parameter sebelum *bootstrap*, $I(B)$ adalah indikator dari kejadian B, dimana 1 jika B benar maka $I(B) = 1$ dan jika B salah maka $I(B) = 0$, $B = Z^* \geq Z_0$ dan R merupakan banyaknya replikasi

Menurut Tu & Zhuo (2003) tahapan lengkap untuk estimasi ATE menggunakan PSSB adalah sebagai berikut.

- a. Menentukan kovariat yang dijadikan sebagai variabel *confounding* berdasarkan teori dan dibuktikan dengan uji dependensi antar variabel.
- b. Menghitung nilai estimasi *propensity score* untuk *full* data dan sampel bootstrap. Estimasi untuk parameter pada sampel bootstrap menggunakan persamaan berikut.

$$\bar{\hat{\beta}} = R^{-1} \sum_{r=1}^R \hat{\beta}^r \quad (1.23)$$

dengan $\hat{\beta}^r$ adalah estimasi parameter replikasi ke- r , $\bar{\hat{\beta}}$ adalah estimasi parameter bootstrap, R adalah banyaknya replikasi.

- c. Membagi subjek menjadi beberapa strata (tergantung peneliti) berdasarkan nilai estimasi *propensity score*. Pembagian strata dilakukan pada *full* data dan sampel bootstrap.
- d. Menguji balance kovariat untuk *full* data dan sampel bootstrap. Pengujian balance dilakukan dengan uji-t dua sampel untuk kovariat bertipe kontinu dan uji-z atau uji chi-square untuk kovariat bertipe kategori.
- e. Menentukan proporsi respon dari kelompok perlakuan \hat{p}_{tk} dan kelompok kontrol \hat{p}_{ck} dalam k subkelas serta jumlah subjek kelompok perlakuan n_t dan kontrol n_c dengan persamaan (2.30).

$$\hat{p}_{tk} = \frac{\sum_{i=1}^{n_k} Y_{tki}}{n_{tk}} \quad ; \quad \hat{p}_{ck} = \frac{\sum_{i=1}^{n_k} Y_{cki}}{n_{ck}}; \quad n_t = \sum_{k=1}^K n_{tk}; \quad n_c = \sum_{k=1}^K n_{ck} \quad (1.24)$$

dengan n_{tk} jumlah subjek kelompok perlakuan kelas ke- k ,
 n_{ck} jumlah subjek kelompok kontrol kelas ke- k ,

$Y_{tk1}, \dots, Y_{tk, n_{tk}}$ respon kelompok perlakuan kelas ke- k ,
 $Y_{ck1}, \dots, Y_{ck, n_{ck}}$ respon kelompok kontrol kelas ke- k , \hat{p}_{tk} dan \hat{p}_{ck}
 adalah proporsi respon dari kelompok perlakuan dan kelompok kontrol kelas ke- k ,

- f. Estimasi Average Treatment Effect (ATE) $\hat{\theta}$ menggunakan persamaan (2.28).

$$\hat{\theta} = \sum_{k=1}^K \frac{n_{tk} + n_{ck}}{n_t + n_c} (\hat{p}_{tk} - \hat{p}_{ck}) \quad (1.25)$$

- g. Menghitung standard error dari ATE dengan persamaan (1.25).

$$SE(\hat{\theta}) = \sqrt{\sum_{k=1}^K \left(\frac{n_{tk} + n_{ck}}{n_t + n_c} \right) \left(\frac{s_{tk}^2}{n_{tk}} + \frac{s_{ck}^2}{n_{ck}} \right)} \quad (1.26)$$

- h. Menentukan sampel *bootstrap* $(\mathbf{X}_i^{(r)}, Z_i^{(r)})$ dimana $i = 1, \dots, n$ yang diambil dari (\mathbf{X}_i, Z_i) dengan pengembalian.
- i. estimasi kembali nilai *propensity score* untuk setiap resample subjek $\hat{e}_i^{(r)}$
- j. Mengulang langkah (c), (e), dan (f) sampai diperoleh estimasi ATE bootstrap $\hat{\theta}^r$
- k. Lakukan langkah (h) sampai (j) sebanyak r replikasi.
- l. Menentukan estimasi *bootstrap* dan standar error menggunakan persamaan (1.27) dan (1.28).

$$\bar{\hat{\theta}} = B^{-1} \sum_{b=1}^B \hat{\theta}^b \quad (1.27)$$

dan

$$SE\left(\bar{\hat{\theta}}\right) = \sqrt{\frac{\sum_{b=1}^B (\hat{\theta}^b - \bar{\hat{\theta}})^2}{(B-1)}} ; b = 1, 2, \dots, B \quad (1.28)$$

m. Menghitung Persen Reduksi Bias (PBR) PSSB menggunakan persamaan (1.11) sampai (1.14).

5. Propensity Score with Marginal Mean Weighting Through Stratification (MMWS)

Pada tahun 2010 Hong memperkenalkan metode *propensity score marginal mean weighting through stratification* (MMWS). Metode *propensity score marginal mean weighting through stratification* merupakan metode yang digunakan untuk menyeimbangkan perlakuan ketika adanya faktor *confounding* (Hong, 2011), dimana unit observasi distratifikasi berdasarkan estimasi *propensity score* dan kemudian dilakukan pembobotan untuk masing-masing nilai *propensity score*. Terdapat beberapa alasan digunakan *propensity score marginal mean weighting through stratification* atau *propensity score* pembobotan melalui stratifikasi yaitu *propensity score* tidak dapat memberikan keseimbangan pada seluruh kovariat dan memasukkan *exposure* dapat menyebabkan adanya hubungan antara beberapa kovariat dan hasil (Ningsih, 2016).

Propensiti score Marginal mean weighting through stratification (MMWS) diperkenalkan sebagai pendekatan yang fleksibel, menggabungkan *propensity score weighting* dan *propensity score stratification* untuk mengatasi ketidakseimbangan karakteristik *intervensi* (Linden, 2014).

Langkah-langkah dalam penerapan MMWS menurut Hong (2011) yaitu:

1. Menentukan estimasi *propensity score*.
2. Membagi sampel kedalam Q strata dengan proporsi yang sama dari setiap kelompok.
3. Melakukan perhitungan *marginal mean weighting through stratification* (MMWS)

$$MMWS = \frac{n_q \times \Pr(Z = t)}{N_{Z=t/c,q}} \quad (1.29)$$

dengan

n_q = Jumlah individu disetiap strata
 $\Pr(Z=t)$ = Probabilitas dari kelompok *treatment*
 $N_{Z=t/c,q}$ = jumlah individu disetiap strata yang diperlakukan sebagai *treatment* dan *kontrol*

4. Melakukan pemeriksaan *balancing* terhadap kovariat
5. Menghitung estimasi efek treatment

Sebelum menentukan nilai *Average treatment effect* (ATE) terlebih dahulu dihitung nilai estimasi *Average treatment effect for the treated* (ATT) untuk kelompok treatment dan kelompok kontrol dengan menggunakan persamaan (1.30) dan (1.31). Selanjutnya menentukan nilai ATE menggunakan persamaan (1.32)

$$E\{\eta(c)\} = \frac{1}{N} \sum_{q=1}^Q \left[\sum_{i \in c,q} \left(\frac{n_{i,q}}{n_{c,q}} \right) \eta_i \right] \quad (1.30)$$

$$E\{\eta(t)\} = \frac{1}{N} \sum_{q=1}^Q \left[\sum_{i \in t,q} \eta_i \right] \quad (1.31)$$

$$ATE = E\{\eta(t)\} - E\{\eta(c)\} \quad (1.32)$$

dengan N adalah jumlah populasi, q adalah jumlah strata, η_i adalah nilai *propensity score* responden ke- i , $n_{t,q}$ adalah jumlah sampel kelompok patuh di setiap sub-strata dalam strata, $n_{c,q}$ adalah jumlah sampel kelompok tidak patuh di setiap sub-strata dalam strata.

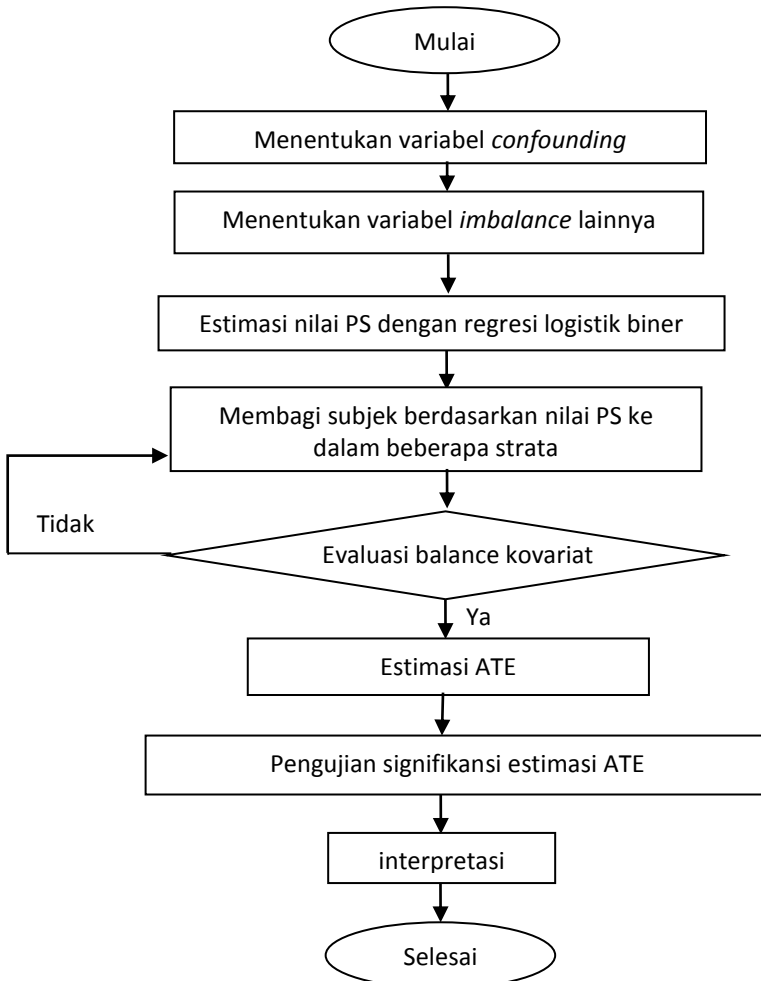
Propensity Score Stratification Using R Language

Software R mempunyai 2 package yang digunakan untuk mengolah data dengan propensity score stratification. Package tersebut adalah *nonrandom package* dan *PSAgraphics package*. Nonrandom package merupakan package yang membahas analisis data secara komprehensif menggunakan propensity score matching dan stratification. Beberapa fungsi pada package ini diimplementasikan mulai dari estimasi nilai propensity score berdasarkan level treatment sampai dengan menguji balance untuk setiap kovariat.

PSA graphics package merupakan *package* yang digunakan untuk menggambar grafik balance setiap kovariat baik kovariat dengan tipe data kategori maupun kontinu. Selain itu package ini dapat digunakan untuk menghitung perbedaan mean setiap strata, estimasi *Average Treatment Effect (ATE)* serta informasi inferensial lainnya tentang treatment. Adapun tahapan propensity score stratification menggunakan program R dibahas pada kasus berikut.

Diberikan data kasus Diabetes Melitus (DM) tipe-2 pada sebuah rumah sakit X dengan responden sebanyak 82 pasien (Akolo, 2016). Data DM terdiri atas 7 variabel yakni 1 variabel

respon berupa Status NDP pasien DM, dan 6 variabel bebas (kovariat) berupa Umur pasien (X1), lama menderita DM (X2), jenis kelamin (X3), Obesitas (X4), Hipertensi (X5), Kadar Asam Urat Serum (X6). Seorang peneliti ingin mengetahui apakah metode propensity score dapat digunakan pada data ini atau tidak. Tahapan propensity score untuk kasus ini digambarkan dalam diagram alir berikut.



Gambar 3.1 Tahapan PSS untuk kasus DM

Penjelasan detail untuk tahapan propensity score stratification (PSS) menggunakan software R adalah sebagai berikut.

1. *Install and load the Nonrandom and PSAgraphics package*

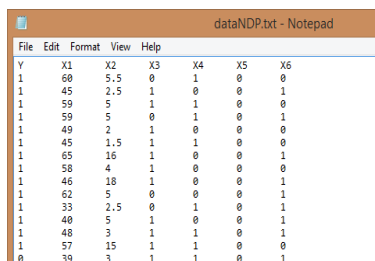
Untuk menggunakan package nonrandom dan PSAgraphics, terlebih dahulu install package dari menu package, jika sudah diinstall selanjutnya load package dengan perintah berikut.

```
## memanggil package yang digunakan
library(nonrandom)
library(PSAgraphics)
```

2. Mempersiapkan data yang digunakan

Untuk memulai tahapan propensity score stratification diperlukan data terkait kasus yang diteliti. Data disimpan dalam notepad (.txt) dengan format variabel sebagai kolom dan observasi (subjek) sebagai baris.

Berikut contoh format data dalam notepad



Y	X1	X2	X3	X4	X5	X6
1	60	5.5	0	1	0	0
1	45	2.5	1	0	0	1
1	59	5	1	1	0	0
1	59	5	0	1	0	1
1	49	2	1	0	0	0
1	45	1.5	1	1	0	0
1	65	16	1	0	0	1
1	58	4	1	0	0	0
1	46	18	1	0	0	1
1	62	5	0	0	0	1
1	33	2.5	0	1	0	1
1	40	5	1	0	0	1
1	48	3	1	1	0	1
1	57	15	1	1	0	0
0	39	3	1	1	0	1

Selanjutnya mendefinisikan kovariat yang bertipe kategori (kecuali X6 yang ditetapkan sebagai variabel *confounding*) sebagai faktor.

```
#read data
> data_NDP<-read.table("D:/lain-lain/dataNDP.txt",
  header = TRUE)
> data_NDP$X3<-factor(data_NDP$X3)
> data_NDP$X4<-factor(data_NDP$X4)
> data_NDP$X5<-factor(data_NDP$X5)
```

3. Memilih kovariat sebagai variabel *confounding*

Pemilihan variabel *confounding* didasarkan pada teori dan bukti empiris tentang hubungan antara kovariat. Berdasarkan teori diabetes mellitus khususnya tentang *Neuropati Diabetic Perifer* (NDP) diketahui bahwa kadar Asam Urat Serum (AUS) pasien DM berhubungan dengan umur pasien, jenis kelamin, obesitas dan lama pasien menderita DM. Selain itu, kadar AUS juga merupakan salah satu faktor resiko kejadian NDP pada pasien DM. oleh karena itu, berdasarkan teori ditetapkan bahwa kadar asam urat serum sebagai variabel *confounding*.

Untuk mendukung teori diatas, diperlukan bukti empiris tentang hubungan antara kadar asam urat serum dengan kovariat lainnya secara statistik. Oleh karena kadar Asam Urat Serum (AUS) merupakan kovariat dengan tipe data kategori, jadi untuk menguji hubungan antara kadar AUS dengan kovariat lain digunakan statistic uji chi-square. Selanjutnya mengetikkan perintah uji chi-square seperti dikolom berikut ini.

```
> C36=chisq.test(X3,X6)
> C46=chisq.test(X4,X6)
> C56=chisq.test(X5,X6)
> pvalue_chisq = cbind(c(C36$p.value, C46$p.value,
  C56$p.value))
> colnames(pvalue_chisq)=c("p-value chi-square")
> rownames(pvalue_chisq)=c("X3*X6", "X4*X6", "X5*X6")
> pvalue_chisq
```


Diperoleh hasil berikut:

```
-----
      p-value chi-square
X3*X6      0.052679799
X4*X6      0.692576697
X5*X6      0.001134814
-----
```

Pvalue_chisq merupakan perintah untuk memberikan output hasil uji chi-square tentang hubungan antara kovariat yang bertipe kategori (X3, X4, X5) dengan variabel *confounding* X6. Hasil pengujian menunjukkan bahwa untuk $\alpha = 5\%$ terdapat 1 kovariat yang mempunyai hubungan dengan variabel *confounding*. Hal ini mengindikasikan bahwa teori yang dijelaskan sebelumnya tentang hubungan antara kovariat dengan variabel *confounding* benar adanya.

4. Mengidentifikasi kovariat bertipe kontinu yang imbalance

Setelah terbukti adanya *confounding* pada data, selanjutnya adalah mengidentifikasi adanya kovariat lain yang kemungkin imbalance dan mengakibatkan bias hasil yang diperoleh. Untuk mengidentifikasi kovariat kontinu yang imbalance digunakan uji t dua sampel. Berikut ini perintah di R untuk uji t dua sampel.

```
##identifikasi kovariat yang imbalance data kontinu
> tx1=t.test(data_NDP$X1[data_NDP$X6 == "0"],
  data_NDP$X1 [data_NDP$X6 == "1"])
> tx2=t.test(data_NDP$X2[data_NDP$X6 == "0"],
  data_NDP$X2 [data_NDP$X6 == "1"])
> summary.identifikasi = cbind(c(tx1$p.value,
  tx2$p.value))
```

```
> colnames(summary.identifikasi)=c("p-value")
> rownames(summary.identifikasi) = c("X1", "X2")
> summary.identifikasi
```

Diperoleh hasil sebagai berikut

```
-----
              p-value
X1 0.554747554
X2 0.004888416
-----
```

Hasil pengujian menunjukkan bahwa kovariat X2 signifikan pada $\alpha = 5\%$ hal ini dapat dilihat dari $p\text{-value} < \alpha$. Hasil ini juga memberikan informasi bahwa kovariat X2 imbalance.

Berdasarkan 2 tahapan ini, diketahui bahwa pada data NDP terdapat *confounding* dan kovariat yang imbalance. Hal ini mengindikasikan bahwa metode propensity score dapat diterapkan pada data ini.

5. Estimasi nilai propensity score

Nilai propensity score pada kasus ini diestimasi menggunakan model regresi logistik. Oleh karena model regresi logistik sama dengan nilai propensity score maka untuk mengetahui nilai propensity, harus diketahui terlebih dahulu parameter modelnya. Parameter model regresi logistik diestimasi dengan metode Maksimum Likelihood. Berikut ini adalah perintah untuk estimasi dan uji signifikansi parameter model regresi logistik.

```
> ps_logit <- glm(X6~X1+X2+X3+X4+X5,data=data_NDP,
  family=binomial("logit"))
> summary(ps_logit)
```

Diperoleh hasil sebagai berikut.

```

Call:
glm(formula = X6 ~ X1 + X2 + X3 + X4 + X5, family =
binomial("logit"), data = data_NDP)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3381 -0.8820  0.5522  0.9024  1.7650

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.12888    1.79863  -0.072  0.9429
X1           -0.01664    0.03343  -0.498  0.6187
X2            0.12815    0.06488   1.975  0.0482 *
X31          -0.76396    0.54489  -1.402  0.1609
X41            0.06482    0.50588   0.128  0.8980
X51            1.14699    0.57556   1.993  0.0463 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
                 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to
be 1)
Null deviance: 113.237  on 81  degrees of freedom
Residual deviance:  95.245  on 76  degrees of
freedom
AIC: 107.24
Number of Fisher Scoring iterations: 4

```

Berdasarkan estimasi diatas, diperoleh bahwa ada 2 kovariat yang signifikan yakni lama pasien menderita DM (X2) dan hipertensi (X5). Dari hasil estimasi juga diperoleh model regresi logistik atau nilai propensity score sebagai berikut.

$$e(\mathbf{x}_i) = \frac{\exp(-0,129 - 0,017X_1 + 0,128X_2 - 0,764X_{31} + 0,065X_{41} + 1,147X_{51})}{1 + \exp(-0,129 - 0,017X_1 + 0,128X_2 - 0,764X_{31} + 0,065X_{41} + 1,147X_{51})} \quad (1.1)$$

Nilai propensity score berkisar antara 0 sampai 1 karena merupakan peluang bersyarat. Berdasarkan persamaan (1.17) diperoleh nilai propensity score berikut

```
# estimasi nilai propensity
NDP.ps <- pscore(data = data_NDP, formula =
X6~X1+X2+X3+X4+X5,
name.pscore = "ps")
NDP.ps$pscore      # perintah print nilai propensity
-----
[1] 0.4115824 0.2106323 0.2370667 0.4001435 0.1897336
[6] 0.2002955 0.5190431 0.2066615 0.6566982 0.372946
[11] 0.4273449 0.2854570 0.2240655 0.5364287 0.251167
[16] 0.2400889 0.2956052 0.1871892 0.2711016 0.1938579
[21] 0.3091197 0.1769140 0.4571176 0.2623075 0.2916421
[26] 0.2379617 0.2493126 0.3266173 0.3951329 0.2652032
[31] 0.3304519 0.1869837 0.1941276 0.8547423 0.4036202
[36] 0.6499749 0.7929604 0.7113100 0.6932462 0.6836200
[41] 0.8817120 0.6222808 0.6350845 0.6836200 0.7436409
[46] 0.4221665 0.8199465 0.8083722 0.8854651 0.6124182
[51] 0.6738146 0.6627512 0.8541277 0.5775472 0.7126948
[56] 0.7664173 0.6269391 0.8414762 0.7731913 0.6015130
[61] 0.7092733 0.6465747 0.8624317 0.7970534 0.7184581
[66] 0.8101773 0.7829718 0.7058308 0.7184581 0.5993209
[71] 0.9349991 0.683620 0.776093 0.623019 0.485423
[76] 0.6195312 0.607106 0.654138 0.626183 0.741405
[81] 0.3539456 0.5493339
```

6. Membagi subjek ke dalam kelas-kelas kecil berdasarkan nilai propensity score

Ada dua cara yang digunakan untuk membagi subjek ke dalam kelas berdasarkan nilai propensity score. Cara yang pertama ada cara manual dan cara kedua dengan memanfaatkan package `nonrandom`.

- Pembagian strata dengan cara manual

Untuk membagi subjek ke dalam kelas (strata) digunakan perintah `cut` pada R. perintah ini menunjukkan bahwa nilai propensity score (`NDP.ps$pscore`) dibagi berdasarkan kuantilnya, dengan syarat `seq(0,1,1/5)` atau nilai propensity dari 0 sampai 1 dibagi 5 kelas sesuai dengan proporsi yang sama ($n/5$). Ada dua keuntungan yang diperoleh pada pembagian strata secara manual, yakni (1) output manual akan menunjukkan secara detail informasi tentang subjek yang masuk di strata 1 sampai strata 5; (2) dapat digunakan pada tahapan estimasi *Average Treatment Effect* (ATE). Perintah yang diketikkan di R untuk pembagian strata manual adalah sebagai berikut.

```
> NDP_s5<- cut(NDP.ps$pscore,
  quantile(NDP.ps$pscore, seq(0, 1, 1/5)),
  include.lowest = TRUE, labels = FALSE)
```

Dan diperoleh hasil pembagian strata seperti pada kolom berikut.

```
[1]  2 1 1 2 1 1 3 1 4 2 2 2 1 3 1 1 2 1 2 1 2 1 2
[24] 1 2 1 1 2 2 1 2 1 1 5 2 4 5 4 4 4 5 3 3 4 5 2
[47] 5 5 5 3 4 4 5 3 4 5 3 5 5 3 4 3 5 5 4 5 5 4 4
[70] 3 5 4 5 3 3 3 3 4 3 4 2 3
```

Output diatas menunjukkan bahwa subjek 1 masuk pada strata 2, subjek 2 dan 3 masuk pada strata 1, dan seterusnya.

- Pembagian strata dengan Package

Package yang digunakan dalam pembagian strata adalah package `nonrandom` dengan perintasan `ps.makestrata`. sama halnya dengan cara manual, perintah `ps.makestrata` ini

menunjukkan bahwa subjek dimasukkan ke dalam strata berdasarkan urutan nilai propensity score, dalam hal ini nilai propensity diurutkan dari yang terkecil (0) sampai yang terbesar (1) berdasarkan proporsi. Jika peneliti hendak membuat 5 strata maka subjek digunakan syarat $\text{seq}(0,1,1/4)$, namun karena dalam penelitian ini digunakan 5 strata maka digunakan syarat $\text{seq}(0,1,1/5)$. Perintah yang diketikkan di console R adalah sebagai berikut.

```
#bagi strata dengan package
> NDP.strata <- ps.makestrata(object = NDP.ps,
  breaks quantile(NDP.ps$pscore, seq(0,1,1/5)),
  stratified.ps = "ps")
> NDP.strata
```

Dan diperoleh hasil sebagai berikut.

```
Stratified by: ps
Strata information:
  Strata bounds      n      n (proportion)
1 [0.177,0.266]    17      0.207
2 (0.266,0.468]   16      0.195
3 (0.468,0.649]   16      0.195
4 (0.649,0.743]   16      0.195
5 (0.743,0.935]   17      0.207
```

Output diatas memberikan informasi bahwa strata 1 memiliki 17 anggota dimana anggotanya merupakan subjek dengan nilai propensity berada pada range 0,177 sampai 0,266, strata 2 memiliki 16 anggota dimana nilai propensity dari anggotanya berada pada range 0,266 sampai 0,468, begitu seterusnya sampai strata 5. Terlihat pula subjek dibagi rata ke 5 strata dengan formulasi $n/5$. Subjek dalam penelitian ini sebanyak 82 orang. Sehingga maka masing-masing strata

memiliki anggota $\frac{82}{5} \approx 16.4$, atau ada strata yang beranggotakan 16 ataupun 17.

Output diatas juga diperoleh dari hasil mengurutkan nilai propensity score dari yang terkecil sampai terbesar, selanjutnya anggota strata 1 adalah subjek dari urutan pertama sampai urutan ke-17, selanjutnya anggota strata 2 adalah subjek dari urutan ke-18 sampai urutan ke 23, begitu seterusnya sampai strata 5.

7. Mengecek balance kovariat untuk setiap strata

Setelah strata sudah dibentuk, maka langkah selanjutnya adalah mengecek balance kovariat untuk setiap strata. Pengujian balance dilakukan berdasarkan tipe data dari kovariat. Untuk kovariat yang bertipe kategori digunakan uji chi-square, sedangkan kovariat yang bertipe kontinu digunakan uji t dua sampel. Pengujian menggunakan perintah `ps.balance` dari package `nonrandom` dengan kovariat yang diuji adalah X1 sampai X5 yang berada pada kolom 2 sampai kolom 6 sesuai format `data_NDP`, sedangkan metode untuk pengujian balance dengan uji t dan uji chi-square menggunakan metode "*classical*". Perintah pengujian balance pada R adalah sebagai berikut.

```
# uji balance
> NDP.balance <- ps.balance(object = NDP.strata,
  sel = c(2:6),method = "classical",
  cat.levels= 2,alpha = 5)
> NDP.balance
```

Diperoleh hasil berikut.

```

Summary of balance check:
                Before: no bal (0) Before: bal (1)
After:no bal (0)                0                0
After:bal (1)                   1                2
Covariates not completely tested:
X3 X5
Detailed balance check (overall):
      X1 X2 X3 X4 X5
Before 1  0  1  1  0
After  1  1 NA  1 NA
Detailed balance check (per stratum):
[p-values from tests (significance level: 0.05)]
      X1    X2    X3    X4    X5
Before 0.555 0.005 0.053 0.693 0.001
-----
Stratum 1 0.054 0.14    NA    1    NA
Stratum 2 0.727 0.057 0.154    1 0.809
Stratum 3 0.545 0.748 0.568    1    1
Stratum 4 0.505 0.839    1    1    1
Stratum 5 0.481 0.764    1    1    NA
-----
Test          t          t chi^2 chi^2 chi^2

```

Output diatas memberikan informasi tentang *p-value* hasil pengujian balance menggunakan uji t dan chi-square. Berdasarkan output diketahui bahwa sebelum strata ada 2 kovariat yang signifikan (*imbalance*) hal ini ditandai dengan nilai *p-value* yang kurang dari $\alpha = 5\%$, kovariat tersebut adalah lama DM (X2) dan hipertensi (X5). Setelah dilakukan stratifikasi terlihat hampir semua kovariat untuk setiap strata memberikan nilai *p-value* yang lebih dari $\alpha = 5\%$. Hal ini mengindikasikan bahwa setelah dilakukan stratifikasi diperoleh

hasil kovariat yang balance di semua strata. Untuk nilai NA pada kovariat X3 di strata 1 serta kovariat X5 di strata 1 dan 5 disebabkan karena ada salah satu kategori yang tidak dimiliki oleh kovariat tersebut. Hal ini dapat lebih jelas dilihat pada grafik balance. Grafik balance kovariat diilustrasikan menggunakan package PSAgraphics dengan perintah **box.psa** untuk data kontinu dan **cat.psa** untuk data kategori. Perintah menggambar grafik di R adalah sebagai berikut.

```
### balance for continous variable__USIA
> box.psa(data_NDP$X1, data_NDP$X6, NDP_s5,
  xlab = "Strata",ylab = "Covariate: usia",
  legend.xy = c(2, 110), balance = F)

### balance for continous variable__LAMA DM
> box.psa(data_NDP$X2, data_NDP$X6, NDP_s5,
  xlab = "Strata",ylab = "Covariate: lama DM",
  legend.xy = c(2, 110), balance = F)

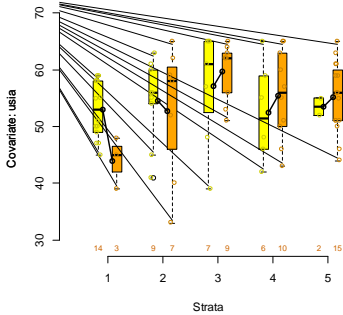
### balance for categorical variable__JK
> cat.psa(data_NDP$X3, data_NDP$X6, NDP_s5,
  xlab = "Strata",
  ylab = "Proportion for 'Jenis Kelamin (X3)'",
  catnames = c("Laki-Laki", "Perempuan"),
  barnames = c("Control", "Treatment"), rtmar = 2)

### balance for categorical variable__Obesitas
> cat.psa(data_NDP$X4, data_NDP$X6, NDP_s5,
  xlab = "Strata",
  ylab = "Proportion for 'Obesitas (X4)'",
  catnames = c("Tidak", "Ya"),
  barnames = c("Control", "Treatment"), rtmar = 2)

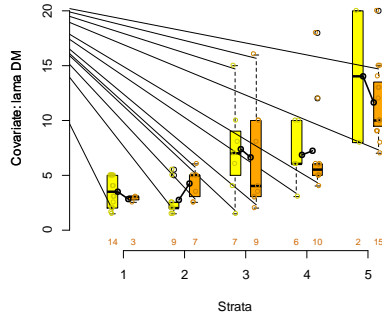
### balance for categorical variable__HIPERTENSI
> cat.psa(data_NDP$X5, data_NDP$X8, NDP_s5,
  xlab = "Strata",
```

```
ylab = "Proportion for 'Hipertensi (X5)'",
catnames = c("Tidak", "Ya"),
barnames = c("Control", "Treatment"), rtmar = 2)
```

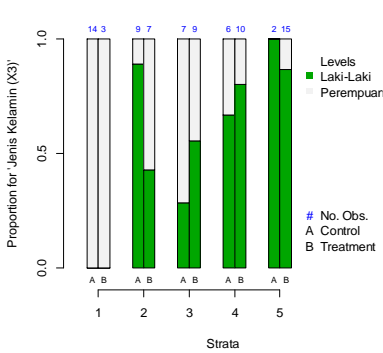
Diperoleh hasil berikut.



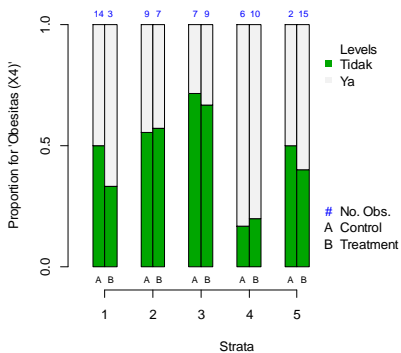
(a)



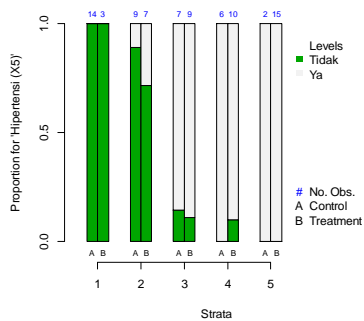
(b)



(c)



(d)



(e)

Gambar 1. Grafik balance kovariat, (a) X1, (b) X2, (c) X3, (d) X4, (e) X5

Berdasarkan gambar diatas terlihat bahwa untuk kovariat jenis kelamin X3 (c) pada strata 1 hanya memiliki anggota perempuan, sedangkan pada untuk kovariat hipertensi X5 (e) pada strata 1 hanya memiliki pasien yang tidak hipertensi dan strata 5 hanya memiliki pasien yang hipertensi. Hal inilah yang menyebabkan nilai p-value NA (*Not Available*).

8. Menghitung mean setiap strata baik untuk kelompok treatment maupun kelompok kontrol, selisih meannya serta estimasi *Average Treatment Effect* (ATE)

Tujuan utama dari metode propensity score adalah menghitung estimasi *Average Treatment Effect* (ATE) yang akurat. Untuk mengestimasi ATE pada R digunakan package PSAGraphics dengan perintah **circ.psa(respon, treatment, strata)**. Berikut ini perintah yang diketikkan di console R untuk estimasi ATE.

```
> circ.psa(Y, X6, NDP_s5, revc = TRUE, ne = 0.1,
  sw = 0.15, inc = 0.17, labcex = 0.9,
  xlab = "Kadar AUS", ylab = "Standard Treatment")
```

Dari perintah diatas, diperoleh output berikut.

```

-----
$summary.strata
  n.0 n.1  means.0  means.1
1   8   9 0.3750000 0.3333333
2   9   7 0.3333333 0.5714286
3  11   5 0.4545455 1.0000000
4   3  13 0.6666667 0.9230769
5   7  10 1.0000000 0.9000000

$wtd.Mn.1
[1] 0.7424238

$wtd.Mn.0
[1] 0.5688747

$ATE
[1] -0.1735491

$se.wtd
[1] 0.1061463

$approx.t
[1] -1.635

$df
[1] 72

$CI.95
[1] -0.38514780 0.03804958

```

Ada beberapa hal yang perlu dibahas dari output diatas, yakni:

1. *Summary Strata*

Summary Strata berisi informasi tentang banyaknya anggota serta mean kelompok kontrol dan treatment pada masing-masing strata.

2. ATE dan se.wtd

Output ATE memberikan nilai estimasi untuk ATE sedangkan se.wtd memberikan nilai estimasi standar error ATE. Berdasarkan output diketahui bahwa estimasi ATE sebesar -0.1735491 dengan standar error 0.1061463.

3. *approx..t* dan *df*

output pada bagian ini memberikan nilai uji t yang diperoleh dari estimasi ATE dibagi dengan standar errornya serta degree of freedom yang diperoleh dari banyaknya respon ($n(Y)$) dikurangi dengan 2 kali banyaknya strata ($2K$). Secara matematis dituliskan sebagai berikut

$$t = \frac{ATE}{se(ATE)} \quad (1.2)$$

$$df = n(Y) - 2(K) \quad (1.3)$$

Hasil output menunjukkan nilai t hitung sebesar -1,635 dengan df 72.

4. CI.95

CI.95 memberikan informasi tentang interval kefidensi ATE pada taraf kepercayaan 95%. Dari output diperoleh nilai ATE berada pada range -0.38514780 dan 0.03804958. hal ini menunjukkan bahwa interval kefidensi ATE memuat 0 sehingga memberikan informasi bahwa estimasi ATE sebesar -0,1735491 tidak berpengaruh signifikan terhadap kejadian NDP atau kadar AUS pasien DM tidak berpengaruh signifikan terhadap kejadian NDP.

Halaman ini sengaja dikosongkan

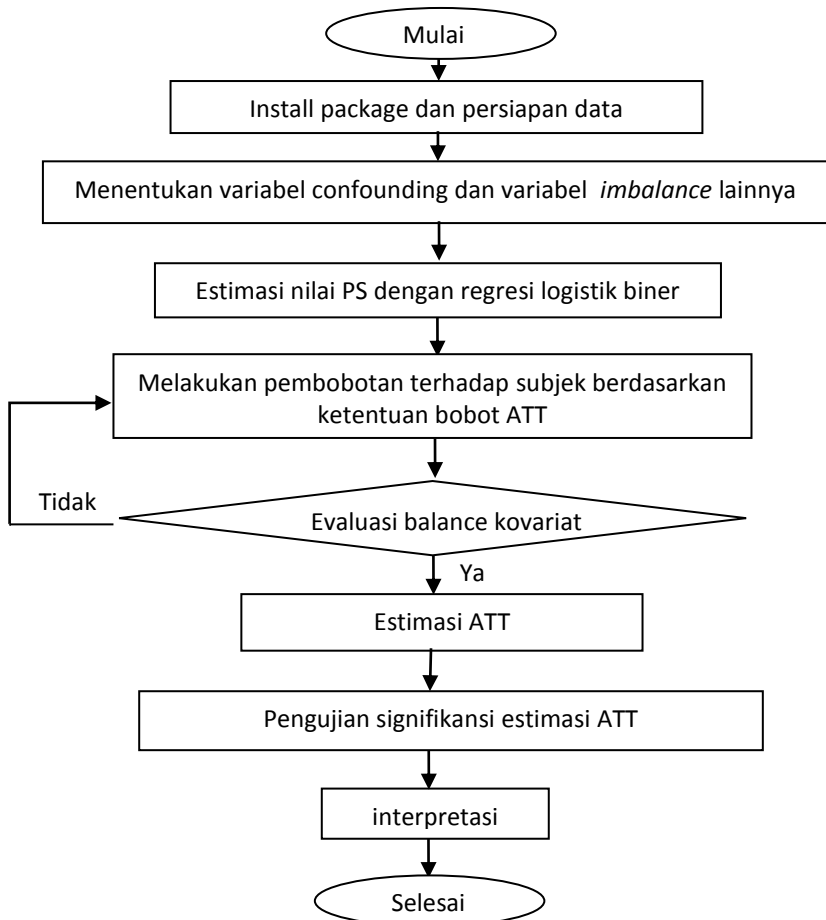
Propensity Score Weighting Using R Language

Untuk mendukung pengolahan data observasi melalui estimasi dan estimasi efek treatment dengan propensity score weighting, software R secara khusus telah menyediakan 1 package yang membahas tentang PSW. Package tersebut adalah *TWANG* (Toolkit for Weighting and Analysis of Nonequivalent Groups) *package*. *Twang* package merupakan package yang bertujuan untuk estimasi propensity score untuk keperluan estimasi efek perlakuan yang lebih akurat, mengecek kualitas dari hasil PSW berdasarkan asumsi balance kovariat, serta menggunakan hasil PSW yang balance untuk menghitung estimasi dari efek treatment.

Seorang peneliti ingin menerapkan metode *Propensity Score Weighting* pada data kasus HIV/AIDS (Islamiyah, 2015). Data HIV yang digunakan terdiri atas 218 responden dengan 14 variabel, yakni 1 variabel respon (Y) yakni kejadian HIV/AIDS, dan 13 variabel prediktor yakni jenis kelamin (X_1), usia (X_2), pendidikan terakhir (X_3), penghasilan (X_4), status nikah (X_5), pasangan tetap (X_6), pasangan tidak tetap (X_7), selalu pakai kondom (X_8), zat yang disuntikkan (X_9), frekuensi suntik (X_{10}), pernah berbagi jarum suntik (X_{11}), selalu pakai

jarum steril (X_{12}), selalu pakai jarum untuk sendiri (X_{13}). Pada kasus ini peneliti ingin mengetahui estimasi dari *Average Treatment Effect on Treated* (ATT).

Sama halnya dengan metode *Propensity Score Stratification* (PSS), metode PSW juga memiliki tahapan yang hampir sama dengan PSS perbedaannya hanya pada bagian pembagian strata untuk PSS dan pembobotan untuk PSW. Berikut ini tahapan PSW untuk kasus diatas menggunakan software R.



Gambar 4.1 Tahapan PSW untuk kasus HIV

1. Install the TWANG Package

Untuk menggunakan package twang, terlebih dahulu install package dari menu package, jika sudah diinstall selanjutnya load package dengan perintah berikut.

```
> library(twang)
```

2. Persiapan Data HIV

Data disimpan dalam notepad (.txt) dengan format variabel sebagai kolom dan observasi (subjek) sebagai baris. Data terdiri atas 14 variabel, dengan 1 variabel respon dan 13 variabel prediktor (kovariat).

Berikut contoh format data dalam notepad

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
0	1	37	2	2	3	1	4	0	2	1	1	2	2
0	1	31	2	2	3	1	2	0	1	2	1	2	2
0	1	30	2	2	2	1	2	0	4	1	1	2	2
0	1	43	2	2	3	3	4	0	1	1	2	2	2
1	1	31	2	1	4	3	4	0	4	3	1	2	2
0	1	37	2	1	4	3	0	0	1	2	2	2	2
0	1	33	2	1	4	1	2	1	1	1	1	2	2
1	2	38	1	2	2	1	1	1	4	1	2	2	2
1	2	28	1	2	2	1	1	0	4	1	2	2	2
0	2	23	1	1	4	1	4	0	4	3	2	2	2
0	1	32	0	1	4	3	4	0	1	1	2	1	2
0	1	29	0	1	4	1	2	0	4	3	2	2	2
1	1	34	1	2	4	3	4	0	1	1	2	2	2
0	1	42	0	1	4	1	2	1	1	2	2	2	1
0	1	32	2	1	4	1	2	0	4	3	1	2	2
0	1	34	2	1	4	3	4	0	1	1	1	2	2
0	1	46	2	1	4	3	4	0	1	1	2	2	2
0	1	35	2	1	4	3	4	0	4	2	1	1	2
0	1	35	2	1	4	3	2	0	4	1	1	1	2
1	1	38	2	1	4	3	4	0	4	1	1	1	2
0	1	31	2	1	4	3	4	0	4	2	1	1	2
0	1	32	2	1	4	3	4	0	4	2	2	2	1
0	1	32	0	1	4	3	4	0	1	1	1	1	2
0	1	35	2	2	2	3	4	0	4	1	1	1	2
1	1	35	2	1	3	3	4	0	1	3	2	2	2
0	1	37	1	1	3	3	4	0	4	2	1	1	2
0	1	39	2	2	3	1	4	0	1	2	1	1	2
0	1	30	2	2	3	1	2	0	4	1	2	2	2
1	1	24	0	1	4	3	4	0	4	3	1	2	2
0	1	35	0	1	4	3	4	0	3	1	1	1	2
0	1	35	2	2	3	1	1	0	4	1	1	1	2
0	1	31	0	1	4	3	4	0	4	2	2	1	2
0	1	32	1	1	4	3	4	0	4	2	2	1	2
0	1	34	2	1	4	3	4	0	4	1	1	1	2
0	1	30	2	1	4	3	4	0	4	1	1	1	2

Tahapan selanjutnya adalah mendefinisikan kovariat yang bertipe kategori sebagai faktor.

```
#read data
> data_NDP<-read.table("D:/lain-lain/dataHIVE.txt",
  header = TRUE)
> data_HIVE$X1<-factor(data_HIVE$X1)
> data_HIVE$X3<-factor(data_HIVE$X3)
> data_HIVE$X4<-factor(data_HIVE$X4)
```

```

> data_HIVF$X5<-factor(data_HIVF$X5)
> data_HIVF$X6<-factor(data_HIVF$X6)
> data_HIVF$X7<-factor(data_HIVF$X7)
> data_HIVF$X8<-factor(data_HIVF$X8)
> data_HIVF$X9<-factor(data_HIVF$X9)
> data_HIVF$X10<-factor(data_HIVF$X10)
> data_HIVF$X11<-factor(data_HIVF$X11)
> data_HIVF$X12<-factor(data_HIVF$X12)
> data_HIVF$X13<-factor(data_HIVF$X13)
> attach(data_HIVF)

```

3. Penentuan Variabel *Confounding*

Syarat awal penentuan variabel *confounding* adalah adanya teori yang menjelaskan hubungan antara kovariat sehingga memungkinkan kovariat tersebut dijadikan sebagai variabel *confounding*. Sesuai dengan teori tentang HIV dijelaskan bahwa ada hubungan antara status nikah seseorang (X5) dan status penggunaan kondom (X8) dengan variabel riwayat penggunaan jarum suntik (X9, X10, X11, X12, X13). Untuk membuktikan teori ini dilakukan pengujian hubungan X5 dan X8 dengan variabel riwayat penggunaan jarum suntik (X9, X10, X11, X12, X13). Pengujian dilakukan menggunakan uji chi-square karena variabel yang ada bertipe kategori. Oleh karena ada dua kovariat yang berpotensi menjadi variabel *confounding*, maka akan dipilih 1 kovariat saja untuk dijadikan sebagai variabel *confounding*. Penentuan kovariatnya didasarkan pada hasil pengujian. Kovariat yang paling banyak memiliki hubungan dengan variabel riwayat jarum suntik (X9, X10, X11, X12, X13) dipilih menjadi variabel *confounding*. Berikut ini perintah yang diketikkan di console R untuk pengujian dependensi menggunakan uji chi-square.

```

#uji dependensi X5 dengan kovariat X9, X10, X11,
  X12, X13
> C59=chisq.test(X5,X9)

```

```

> C510=chisq.test(X5,X10)
> C511=chisq.test(X5,X11)
> C512=chisq.test(X5,X12)
> C513=chisq.test(X5,X13)

## uji dependensi X8 dengan kovariat X9,X10,
  X11,X12,X13
> C89=chisq.test(X8,X9)
> C810=chisq.test(X8,X10)
> C811=chisq.test(X8,X11)
> C812=chisq.test(X8,X12)
> C813=chisq.test(X8,X13)
> pvalue_chisq = cbind(c(C59$p.value,C510$p.value,
  C511$p.value,C512$p.value,C513$p.value,
  C89$p.value,C810$p.value, C811$p.value,
  C812$p.value,C813$p.value))
> colnames(pvalue_chisq)=c("p-value chi-square")
> rownames(pvalue_chisq)=c("X5*X9","X5*X10",
  "X5*X11","X5*X12","X5*X13","X8*X9","X8*X10",
  "X8*X11","X8*X12","X8*X13")
> pvalue_chisq

```

Diperoleh hasil berikut

	p-value chi-square
X5*X9	0.0003384248
X5*X10	0.3745694882
X5*X11	0.2371141366
X5*X12	0.3698778040
X5*X13	0.0961835536
X8*X9	1.137213e-01
X8*X10	7.420376e-01
X8*X11	8.037329e-02
X8*X12	8.897221e-02
X8*X13	8.558044e-06

Output diatas menunjukkan bahwa pada pengujian hubungan X5 dengan X9 sampai X13 diketahui bahwa untuk $\alpha = 10\%$ ada 2 kovariat yang signifikan, yakni X5*X9 dan X5*X13. Hal ini memberikan gambaran bahwa ada hubungan antara kovariat X5 dengan X9 dan X13. Untuk pengujian hubungan X8 dengan X9 sampai X13 diketahui bahwa untuk $\alpha = 10\%$ ada 3 kovariat yang signifikan, yakni X8*X11, X8*X12, dan X8*X13. Hal ini mengindikasikan bahwa ada hubungan antara kovariat X8 dengan X11, X12 dan X13.

Berdasarkan hasil pengujian diatas disimpulkan bahwa kovariat X8 yang paling banyak berhubungan dengan variabel riwayat jarum suntik (X9 sampai X13). Oleh karena itu, kovariat X8 ditetapkan sebagai variabel *confounding*.

4. Identifikasi Kovariat Kontinu yang *Imbalance*

Setelah terbukti adanya *confounding* pada data, selanjutnya adalah mengidentifikasi adanya kovariat lain yang kemungkinan *imbalance* pada data. Untuk mengidentifikasi kovariat kontinu yang *imbalance* digunakan uji t dua sampel. Berikut ini perintah di R untuk uji t dua sampel.

```
> tx2=t.test(data_HIVF$X2[data_HIVF$X8 == "0"],
  data_HIVF$X2 [data_HIVF$X8 == "1"])
> summary.identifikasi = cbind(c(tx2$p.value))
> colnames(summary.identifikasi)=c("p-value")
> rownames(summary.identifikasi) = c("X2")
> summary.identifikasi
```

Diperoleh hasil sebagai berikut

	p-value
X2	0.982

Hasil pengujian menunjukkan bahwa kovariat X2 signifikan pada $\alpha = 5\%$ hal ini dapat dilihat dari p-value $> \alpha$. Hasil ini juga

memberikan informasi bahwa tidak ada kovariat kontinu yang *imbalance*.

5. Estimasi Nilai Propensity Score

Setelah syarat perlu dari propensity score terpenuhi dalam hal ini terbukti adanya variabel *confounding* pada data, maka dapat dilanjutkan dengan estimasi nilai propensity score. Sebelum melakukan estimasi nilai propensity, hendaknya terlebih dahulu memperbaiki data yang ada. Hal ini disebabkan karena kovariat yang dijadikan *confounding* tidak bisa didefinisikan sebagai faktor. Untuk mengubah definisi *factor* menjadi *unfactor* digunakan *package varhandle* dengan perintah **unfactor**. Selanjutnya estimasi nilai propensity dilakukan menggunakan model regresi logistik dengan metode *maximum likelihood*. Estimasi nilai propensity menggunakan perintah **glm** seperti yang dituliskan berikut ini.

```
##estimasi nilai propensity
> library(varhandle)
> data_HIVF$X8<- unfactor(data_HIVF$X8)
> ps_logitHIV<-glm(X8~X1+X2+X3+X4+X5+X7+X9+X10+
  X11+X12+X13,data=data_HIVF,
  family=binomial("logit"))
> summary(ps_logitHIV)
```

Diperoleh hasil estimasi parameter untuk nilai propensity sebagai berikut.

```
Call:  glm(formula = X8 ~ X1 + X2 + X3 + X4 + X5 +
  X7 + X9 + X10 + X11 + X12 + X13, family =
  binomial("logit"), data = data_HIVF)
```

```
Coefficients:
```

(Intercept)	X12	X2	X31	X32
-8.46931	3.75663	-0.04228	-1.28749	-0.35320
X33	X42	X52	X53	X54
3.13975	-0.85874	13.54974	11.31318	11.85723
X62	X63	X72	X73	X74
1.36804	0.81839	-0.68645	-23.02813	-2.69305
X92	X93	X94	X95	X102
-17.71269	-16.79060	-0.96007	0.53554	0.41031
X103	X112	X122	X132	
0.22755	-2.19210	-2.19210	-0.42634	
Degrees of Freedom: 217 Total (i.e. Null); 196 Residual				
Null Deviance: 119.4				
Residual Deviance: 75.7		AIC: 123.7		

Dari hasil estimasi juga diperoleh model regresi logistik atau nilai propensity score sebagai berikut.

$$e(x_i) = \frac{\exp(K_1)}{1 + \exp(K_1)} \quad (1.1)$$

dengan

$$\begin{aligned}
 K_1 = & -8,469 + 3,757x_{12} - 0,042x_2 - 1,287x_{3,1} - 0,353x_{3,2} + 3,140x_{3,3} \\
 & - 0,859x_{4,2} + 13,550x_{5,2} + 11,313x_{5,3} + 11,857x_{5,4} + 1,368x_{6,2} \\
 & + 0,818x_{6,3} - 0,686x_{7,2} - 23,028x_{7,3} - 2,693x_{7,4} - 17,713x_{9,2} \\
 & - 16,791x_{9,3} - 0,960x_{9,4} + 0,536x_{9,5} + 0,410x_{10,2} + 0,228x_{10,3} \\
 & - 2,192x_{11,2} - 0,426x_{12,2} - 2,722x_{13,2}
 \end{aligned}$$

Nilai propensity score berkisar antara 0 sampai 1 karena merupakan peluang bersyarat. Berikut ini perintah untuk mengetahui nilai propensity score menggunakan perintah **predict**.

```
> Nilai_ps<- predict(ps_logitHIV,type="response")
> Nilai_ps
```

Berdasarkan persamaan (1.23) diperoleh nilai propensity score berikut

	1	2	3	4
	9.667285e-11	4.299345e-02	1.002628e-01	6.104433e-04
	5	6	7	8
	1.744677e-02	3.190900e-03	1.002145e-01	3.115151e-01
	9	10	11	12
	3.115151e-01	2.021489e-02	8.549260e-03	9.980065e-03
	13	14	15	16
	3.087568e-11	2.171509e-01	6.198844e-03	3.151335e-02
	⋮			
	⋮			
	211	212	213	214
	8.962186e-02	6.787642e-04	1.124964e-03	1.947499e-03
	215	216	217	218
	1.040787e-02	2.220446e-16	5.079785e-03	5.399737e-02

6. Pembobotan terhadap Subjek untuk Setiap Kovariat

Pembobotan yang dilakukan terhadap subjek bergantung pada estimasi treatment yang diinginkan. Berdasarkan kasus yang dijelaskan sebelumnya bahwa peneliti ingin mengetahui estimasi dari *Average Treatment Effect on Treated* (ATT), sehingga bobot subjek disesuaikan dengan kriteria untuk estimasi ATT. Pada teori sebelumnya diketahui bahwa untuk estimasi ATT, bobot yang diberikan pada subjek kelompok treatment adalah $w=1$ dan bobot untuk kelompok kontrol

adalah $w = \frac{e(\mathbf{x}_i)}{(1 - e(\mathbf{x}_i))}$. Hal ini berarti subjek kelompok

treatment nilainya tidak berubah karena diberi bobot 1, sedangkan subjek kelompok kontrol nilai berubah sebesar w kali. Berikut ini merupakan perintah pembobotan di R untuk kelompok kontrol.

```
> data_HIVF$w.logit <- rep(1,nrow(data_HIVF))
> data_HIVF$w.logit[data_HIVF$X8==0] <-
  exp(predict(ps_logitHIV, subset(data_HIVF, X8==0)))
```

Diperoleh hasil pembobotan untuk subjek pada kelompok kontrol adalah sebagai berikut.

```
[1] 9.667285e-11 4.492492e-02 1.114356e-01
[4] 6.108162e-04 1.775657e-02 3.201114e-03
[7] 4.524647e-01 2.063196e-02 8.622980e-03
[10] 1.008067e-02 3.087568e-11 6.237509e-03
[13] 3.253875e-02 2.188013e-03 2.757022e-02
[16] 1.519430e-02 1.611240e-02 1.440330e-02
[19] 3.469932e-02 7.720996e-02 4.210597e-02
:
:
[199] 2.448311e-18 5.105721e-03 5.707951e-02
```

7. Mengecek Balance Kovariat

Setelah dilakukan pembobotan, maka langkah selanjutnya adalah mengecek balance setiap kovariat. Untuk mengecek balance kovariat digunakan perintah **`dx.wts(data yang sudah diboboti, data awal, treatment, perm.test.iters=0, estimand="ATE" atau "ATT")`**. Untuk lebih jelasnya perintah cek balance kovariat dapat dilihat tabel berikut.

```
> bal.logit <- dx.wts(data_HIVF$w.logit,
  data=data_HIVF, treat.var="X8",
  perm.test.iters=0, estimand = "ATT")
> pretty.tab <- bal.table(bal.logit)[[2]][,]
```



```

c("tx.mn", "ct.mn", "ks") ]
> pretty.tab <- cbind(pretty.tab,
  bal.table(bal.logit)[[1]]$ct.mn)
> names(pretty.tab) <- c("E(Y1|t=1)", "E(Y0|t=1)",
  "KS", "E(Y0|t=0)")
> xtable(pretty.tab,
  caption = "Logistic regression estimates of the
  propensity scores", label = "tab:balancelogit",
  digits = c(0, 2, 2, 2, 2),
  align=c("l", "r", "r", "r", "r"))

```

Diperoleh hasil pengujian balance dengan uji Kolmogorv-Smirnov sebagai berikut.

	$E(Y1 \$t=1)$	$E(Y0 \$t=1)$	KS	$E(Y0 \$t=0)$
X1:1	0.94	0.94	0.00	0.97
X1:2	0.06	0.06	0.00	0.03
X2	32.71	32.90	0.15	32.68
X3:0	0.18	0.17	0.00	0.21
X3:1	0.06	0.05	0.01	0.11
X3:2	0.65	0.64	0.00	0.65
X3:3	0.12	0.13	0.01	0.03
X4:1	0.65	0.68	0.03	0.66
X4:2	0.35	0.32	0.03	0.34
X5:1	0.00	0.00	0.00	0.01
X5:2	0.29	0.28	0.02	0.12
X5:3	0.12	0.10	0.02	0.15
X5:4	0.59	0.62	0.04	0.71
X6:1	0.29	0.32	0.02	0.37
X6:2	0.06	0.02	0.04	0.02
X6:3	0.65	0.66	0.01	0.61
X7:1	0.06	0.06	0.00	0.01
X7:2	0.29	0.35	0.06	0.31
X7:3	0.00	0.00	0.00	0.02
X7:4	0.65	0.59	0.06	0.65

X9:1	0.53	0.57	0.04	0.30
X9:2	0.00	0.00	0.00	0.03
X9:3	0.00	0.00	0.00	0.04
X9:4	0.41	0.41	0.00	0.62
X9:5	0.06	0.02	0.04	0.01
X10:1	0.59	0.66	0.07	0.62
X10:2	0.29	0.25	0.05	0.22
X10:3	0.12	0.09	0.02	0.16
X11:1	0.76	0.73	0.03	0.51
X11:2	0.23	0.27	0.03	0.49
X12:1	0.65	0.66	0.01	0.40
X12:2	0.35	0.34	0.01	0.60
X13:1	0.59	0.52	0.07	0.13
X13:2	0.41	0.48	0.07	0.87
w.logit	1.00	0.65	0.68	0.08

Berdasarkan output diatas diketahui bahwa rata-rata treatment dan kontrol untuk semua kovariat tidak jauh berbeda. Hal ini juga didukung oleh hasil pengujian balance dengan uji Kolmogorov-Smirnov yang menunjukkan bahwa treatment dan kontrol berasal dari distribusi yang sama atau balance untuk semua kovariat. Hasil ini dibuktikan dengan nilai KS_{hitung} pada kolom KS lebih dari nilai $KS_{tabel} = 0,3435$.

8. Estimasi ATT

Setelah syarat balance kovariat telah terpenuhi, maka tahapan dapat dilanjutkan ke estimasi ATT. Perintah yang digunakan untuk estimasi ATT adalah sebagai berikut.

```
> design.logit <- svydesign(ids=~1,
  weights=~w.logit, data=data_HIVF)
> glm6 <- svyglm(Y ~ X8, design=design.logit)
> summary(glm6)
```

Diperoleh hasil sebagai berikut.

```

Call:
svyglm(formula = Y ~ X8, design = design.logit)
Survey design:
svydesign(ids = ~1, weights = ~w.logit, data =
data_HIVF)

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.355643    0.092212   3.857 0.000152
***
X8            -0.002701    0.148320  -0.018 0.985485
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
                 '.' 0.1 ' ' 1
(Dispersion parameter for gaussian family taken to
be 0.2298134)
Number of Fisher Scoring iterations: 2

```

Berdasarkan output diatas diketahui bahwa estimasi ATE adalah sebesar -0,0027 adengan standar error sebesar 0,1483. Akan tetapi nilai *p-value* = 0,985 yang lebih dari $\alpha = 10\%$ menunjukkan bahwa treatment tidak berpengaruh terhadap kejadian HIV/AIDS.

Halaman ini sengaja dikosongkan

Apply of Propensity Score with Marginal Mean Weighting Through Stratification (MMWS) using R

Metode propensity score dengan MMWS ini tidak memiliki package khusus yang bisa digunakan di software R. Hal ini berbeda dengan metode lainnya yang dibahas sebelumnya. Oleh karena itu, untuk mengaplikasikan metode propensity score dengan MMWS diperlukan syntax yang mengurai pengerjaan dengan cara manual berdasarkan tahapan propensity score dengan MMWS. Untuk lebih jelasnya dapat dilihat pada kasus berikut.

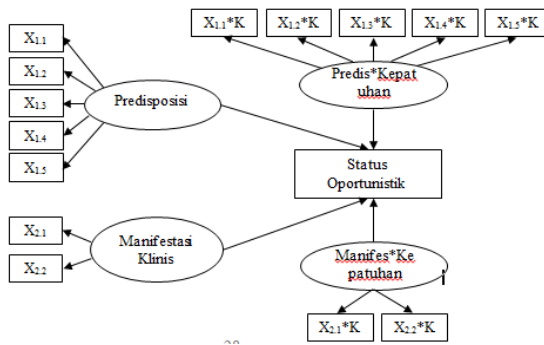
Diberikan data observasi tentang kasus HIV/AIDS. Data berupa rekam medis pasien HIV/AIDS yang berjumlah 90 pasien. Data pada kasus ini terdiri atas 3 variabel laten, yakni 1 variabel endogen Y_1 (infeksi oportunistik) dan 2 variabel eksogen X_1 (predisposisi) dan X_2 (manifestasi klinis). Oleh karena data ini merupakan data observasi maka ada kemungkinan terdapat variabel *confounding* yang mengakibatkan estimasi efek treatment tidak akurat (Ningsih, 2016). Untuk menanggulangi masalah *confounding* ini peneliti menerapkan metode *Propensity Score With Marginal Mean Weighting Through Stratification* (MMWS). Metode propensity score dengan MMWS ini digunakan khusus untuk data yang memiliki variabel laten didalamnya. Berikut ini dijelaskan 2 variabel laten endogen yang digunakan dalam penelitian.

Tabel 5.1. Variabel laten dan indikator

Variabel Laten (Konstruk)	Variabel manifes (indikator)
Predisposisi (X ₁)	X _{1.1} (Usia)
	X _{1.2} (Jenis Kelamin)
	X _{1.3} (Tingkat Pendidikan)
	X _{1.4} (Pekerjaan)
	X _{1.5} (Status Perkawinan)
Manifestasi Klinis (X ₂)	X _{2.1} (Kadar CD4)
	X _{2.2} (Stadium Klinis)

Oleh karena variabel yang digunakan bersifat laten dan tidak terukur, maka untuk membuktikan adanya variabel *confounding* pada data digunakan model yang secara khusus membahas tentang variabel laten, yakni model *Structural Equation Modelling* (SEM). Model SEM yang digunakan adalah model SEM-PLS. berikut ini tahapan propensity score MMWS yang diawali dengan pengujian efek *confounding* dengan model SEM-PLS.

1. Melakukan analisis pada efek *confounding* dengan model SEM-PLS.
 - a. Pengembangan model berbasis konsep dan teori. Berdasarkan teori diketahui bahwa kepatuhan memiliki hubungan dengan predisposisi dalam mempengaruhi Infeksi oportunistik. Demikian juga kepatuhan memiliki hubungan dengan manifestasi klinis dalam mempengaruhi infeksi oportunistik.
 - b. Mengkonstruksi diagram path untuk model persamaan struktural dengan hubungan kepatuhan.



Gambar 5.1 Diagram Path Infeksi Oportunistik dengan Variabel Coufounding

- Mengkonversi diagram path pada gambar 5.1i kedalam bentuk persamaan matematis baik model pengukuran maupun model struktural sebagai berikut.

$$IO = \gamma_1 \text{ Predis} + \gamma_2 \text{ Manifes} + \omega_1 \text{ Predis*Ke patuhan} + \omega_2 \text{ Manifes*Ke patuhan} \quad (1.1)$$

- Menentukan efek variabel *confounding*

Setelah mendapatkan model diagram *path*, selanjutnya dilakukan pengujian signifikansi variabel *confounding* pada model (5.1) dengan hipotesis sebagai berikut:

H_3 :Kepatuhan memiliki hubungan dengan predisposisi dalam mempengaruhi infeksi oportunistik

H_4 :Kepatuhan memiliki hubungan dengan Manifestasi klinis dalam mempengaruhi infeksi oportunistik

Daerah kritis : tolak H_0 jika $t_{stat} > t_{tabel}$ atau $p\text{-value} < \alpha$.

Dengan taraf signifikansi $\alpha = 10\%$ dan *software PLS* diperoleh hasil pengujian variabel *confounding* berikut.

Tabel 5.2 Signifikansi model struktural dengan variabel *confounding*

Hubungan Kausalitas	Koefisien Parameter	Standard error	t-stat
Predisposisi -> IO	0.315	0.071	4.415
Manifestasi Klinis -> IO	0.636	0.059	10.833
Predis*Kepatuhan -> IO	-0.158	0.078	2.019
Manifes* Kepatuhan -> IO	0.123	0.068	1.805

Berdasarkan Tabel 5.2 dapat dilihat bahwa koefisien parameter jalur yang diperoleh dari hubungan antara predisposisi dengan kepatuhan terapi ARV terhadap infeksi oportunistik sebesar -0,158 dengan nilai t-hitung = 2,019 > $t_{\frac{\alpha}{2}, 90-2} = 1,64$ pada taraf signifikansi $\alpha = 10\%$. Hasil ini

mengindikasikan bahwa faktor *confounding* (kepatuhan terapi ARV) memiliki hubungan yang signifikan dengan predisposisi dalam mempengaruhi infeksi oportunistik. Demikian juga koefisien parameter jalur yang diperoleh dari hubungan antara manifestasi klinis dengan kepatuhan terapi ARV terhadap infeksi oportunistik sebesar 0,123 dimana faktor *confounding* (kepatuhan terapi ARV) memiliki hubungan yang signifikan dengan manifestasi klinis dalam mempengaruhi infeksi oportunistik. Hal ini ditunjukkan dengan nilai t-hitung = 1,805 > $t_{\frac{\alpha}{2}, 90-2} = 1,64$ pada taraf signifikansi $\alpha = 10\%$.

4. Mempersiapkan data

Setelah ditentukan variabel *confounding*, selanjutnya menganalisis propensity score menggunakan software R.

Langkah awal yang dilakukan adalah mempersiapkan data yang digunakan dan mendefinisikan variabel penelitian. Variabel yang dijadikan kovariat pada tahapan ini adalah nilai skor faktor masing-masing kovariat yang diperoleh pada tahapan SEM-PLS sebelumnya. Nilai skor faktor untuk kovariat predisposisi (ξ_1) dan kovariat manifestasi klinis (ξ_1) adalah sebagai berikut.

Tabel 5.3 Nilai *Score Factor*

Predis posisi	Manifestasi Klinis	Lanjutan		Lanjutan	
		Predis posisi	Manifestasi Klinis	Predis posisi	Manifestasi Klinis
0.142	0.025	-0.597	-0.539	0.156	0.768
1.975	0.204	-1.260	1.332	-0.521	-0.539
-1.260	-1.282	1.975	-1.846	0.232	0.768
1.113	0.025	-1.260	0.025	-0.597	-0.539
-0.507	0.768	-0.521	0.025	-1.260	-0.539
-0.507	0.768	1.004	1.332	-1.260	0.768
⋮	⋮	⋮	⋮	⋮	⋮
0.142	1.332	-0.507	0.025	-1.260	-1.846
-0.597	-1.846	1.975	-0.539	-1.260	1.332
-0.507	0.768	-0.507	1.332	1.648	1.332

Nilai skor faktor diatas disimpan di data D dalam bentuk file txt, dengan nama file HIV.txt. contoh format datanya dapat dilihat pada gambar berikut.

File	Edit	Format	View	Help
Predisposisi			Manifestasi Klinis	
0.142			0.025	
1.975			0.204	
-1.260			-1.282	
1.113			0.025	
-0.507			0.768	
-0.507			0.768	
0.142			-1.046	
-1.260			-0.539	
0.909			0.768	
1.004			0.025	
-0.507			1.332	
-1.757			0.204	
1.236			-1.046	
1.236			-0.539	
1.004			0.025	
-1.260			-0.539	
-1.260			-1.046	
-0.507			-0.539	
1.004			1.332	
1.004			0.025	
-0.398			-1.046	
0.895			0.025	
-0.521			-1.046	
1.004			1.332	
1.975			0.204	
-0.507			-0.539	
-1.260			0.768	
0.142			1.332	
-0.507			-1.046	
-0.487			0.768	

Berikut ini syntax manual yang digunakan untuk mempersiapkan data.

```

> rm(list=ls()) #Restart
> Propensity = read.table("D://HIV.txt", header=TRUE)
#Lokasi data
> Psi1 = Propensity$Psi1 #Variabel Latent 1
> Psi2 = Propensity$Psi2 #Variabel Latent 2
> G1 = Propensity$G1 #Variabel Gamma 1
> G2 = Propensity$G2 #Variabel Gamma 2
> KPH = Propensity$KPH #Kepatuhan
> Nmr = Propensity$No #Nomor Pasien
> k=3 #Jumlah strata yang ingin dibentuk
> n=length(Nmr) #Jumlah data

> Psi1 = matrix(Psi1,n,1) #Matrix Variabel Latent 1
> Psi2 = matrix(Psi2,n,1) #Matrix Variabel Latent 2
> G1 = matrix(G1,n,1) #Matrix Variabel Gamma 1
> G2 = matrix(G2,n,1) #Matrix Variabel Gamma 2
> KPH = matrix(KPH,n,1) #Matrix Kepatuhan

```

5. Menentukan estimasi *propensity score*

Setelah diperoleh estimasi koefisien jalur untuk masing-masing variabel laten dengan menggunakan SEM-PLS, maka langkah selanjutnya adalah mengestimasi *propensity score* dengan menggunakan model dengan nilai koefisien parameter seperti pada Tabel 5.4.

Tabel 5.4 Koefisien parameter

Variabel Laten	Koefisien Parameter
Predisposisi	0,269
Manifestasi Klinis	0,688

Berdasarkan estimasi koefisien parameter, diperoleh nilai propensity score sebagai berikut.

$$e(\eta_i) = \frac{\exp(0,269\xi_{i1} + 0,688\xi_{i2})}{1 + \exp(0,269\xi_{i1} + 0,688\xi_{i2})} \quad (1.2)$$

Untuk menghitung nilai *propensity score* secara manual diberikan ilustrasi berikut, misalkan untuk responden 1 diketahui nilai skor faktor $\xi_1 = 0,142$ dan $\xi_2 = 0,025$ (lihat tabel 5.3) maka nilai propensity score yang diperoleh adalah 0,5138 seperti pengerjaan dibawah ini. Begitu seterusnya sampai diperoleh nilai propensity score untuk semua responden.

$$e(\eta_1) = \frac{\exp(0,269 \times 0,142 + 0,688 \times 0,025)}{1 + \exp(0,269 \times 0,142 + 0,688 \times 0,025)}$$

$$e(\eta_1) = \frac{1,056961}{2,056961} = 0,513846$$

Untuk menghitung nilai propensity score menggunakan software R dapat digunakan perintah berikut.

```
> GP1 = matrix(G1*Psi1,n,1) #Matrix Perkalian Variabel
  Latent 1 dengan Variabel Gamma 1
> GP2 = matrix(G2*Psi2,n,1) #Matrix Perkalian Variabel
  Latent 2 dengan Variabel Gamma 2
> GP = GP1+GP2 #Matrix Penjumlahan dari Matrix Gamma dan
  Matrix Psi
> expGP = exp(GP) #Pembilang dari Matrix Nilai Estimasi
  Propensity Score atau Exponen dari Matrix GP
> exp2GP = 1+exp(GP) #Penyebut dari Matrix Nilai
```

```

Estimasi Propensity Score atau 1+Exponen Matrix GP
> e = expGP / exp2GP #Matriks Nilai Estimasi Propensity
Score
> e_sort = matrix(e[order(-e),]) #Nilai Estimasi
Propensity Score diurutkan

```

Nilai propensity score yang diperoleh untuk 90 responden dapat dilihat pada Tabel 5.5.

Tabel 5.5 Nilai *Propensity score*

Respon- den	$e(\eta_i)$	Lanjutan		Lanjutan	
		Respon- den	$e(\eta_i)$	Respon- den	$e(\eta_i)$
1	0,513846	31	0,370185	61	0,638844
2	0,661867	32	0,640486	62	0,374964
3	0,227768	33	0,323273	63	0,643547
⋮	⋮	⋮	⋮	⋮	⋮
29	0,192999	59	0,540025	89	0,640486
30	0,596764	60	0,685687	90	0,795716

6. Membagi sampel kedalam Q strata berdasarkan nilai *propensity score* yang telah diurutkan.

Nilai propensity score yang diperoleh pada tahapan sebelumnya akan digunakan untuk membagi sampel ke dalam strata. Pembagian strata di R dilakukan berdasarkan nilai kuantil dari propensity score yang telah diurutkan. Banyaknya strata yang akan dibuat adalah 3 strata. Berikut syntax untuk pembagian strata di R.

```

> PropensityDF = data.frame(Nmr=c(Nmr),
  Psi1=c(Psi1), Psi2=c(Psi2), G1=c(G1), G2=c(G2),
  e=c(e), Kepatuhan=c(KPH))
> Strata = quantile(PropensityDF$e,
  prob = seq(from = 0, to = 1, by = (1/k)),

```

```

na.rm = TRUE)
> PropensityDF$Strata = cut(PropensityDF$e,
  breaks= Strata, labels=c('3', '2', '1'),
  include.lowest=TRUE)

```

Hasil pembagian strata menggunakan syntax diatas ditunjukkan oleh Tabel 5.6.

Tabel 5.6 Pembagian Jumlah Sub-Strata berdasarkan nilai *propensity*

Strata	Sub-Strata	Patuh	Tidak Patuh	Jumlah responden
	1	7	23	30
3	2	15	14	29
	3	25	6	31

Tabel 5.6 menunjukkan bahwa strata 1, 2 dan 3 memiliki jumlah responden yang hampir sama karena strata dibagi rata menjadi 3 strata dengan distribusi proporsi yang sama.

7. Menghitung Nilai MMWS untuk setiap kelompok patuh dan kelompok tidak patuh pada setiap strata, untuk kelompok patuh menggunakan persamaan (3.4) dan kelompok tidak patuh menggunakan persamaan (3.5)

$$MMWS = \frac{n_q \times \Pr(Z = t)}{n_{Z=t,q}} \quad (1.3)$$

$$MMWS = \frac{n_q \times \Pr(Z = c)}{n_{Z=c,q}} \quad (1.4)$$

Nilai MMWS merupakan nilai yang penting untuk diketahui karena setiap responden kelompok treatment dan kontrol diboboti dengan nilai MMWS. Berikut ini syntax R untuk

menghitung nilai MMWS setiap strata dan membuat strata baru berdasarkan nilai responden yang sudah diboboti.

```
#Strata dan MMWS
> StrA = length(PropensityDF$Nmr[PropensityDF$Strata==1
  & PropensityDF$Kepatuhan==1])
> StrB = length(PropensityDF$Nmr[PropensityDF$Strata==3\2
  & PropensityDF$Kepatuhan==1])
> StrC = length(PropensityDF$Nmr[PropensityDF$Strata==3
  & PropensityDF$Kepatuhan==1])
> StrG = length(PropensityDF$Nmr[PropensityDF$Strata==1
  & PropensityDF$Kepatuhan==2])
> StrH = length(PropensityDF$Nmr[PropensityDF$Strata==2
  & PropensityDF$Kepatuhan==2])
> StrI = length(PropensityDF$Nmr[PropensityDF$Strata==3
  & PropensityDF$Kepatuhan==2])

> StrR =length(PropensityDF$Nmr[PropensityDF$Strata==1])
> StrS =length(PropensityDF$Nmr[PropensityDF$Strata==2])
> StrT =length(PropensityDF$Nmr[PropensityDF$Strata==3])

> StrP=length(PropensityDF$Nmr[PropensityDF$Kepatuhan==1])
> StrQ=length(PropensityDF$Nmr[PropensityDF$Kepatuhan==2])
> StrZ = length(PropensityDF$Nmr)

#Tabel Strata
> Str = data.frame(Strata=c('1','2','3','Total'),
  Patuh=c(StrA,StrB,StrC,StrP),
  TidakPatuh=c(StrG,StrH,StrI,StrQ),
  Total=c(StrR,StrS,StrT,StrZ))

> MMWSA = (StrR * (StrP/StrZ))/StrA #Nilai MMWS Strata 1
  yang Patuh
> MMWSB = (StrS * (StrP/StrZ))/StrB #Nilai MMWS Strata 2
  yang Patuh
> MMWSC = (StrT * (StrP/StrZ))/StrC #Nilai MMWS Strata 3
  yang Patuh

> MMWSG = (StrR * (StrQ/StrZ))/StrG #Nilai MMWS Strata 1
  yang Tidak Patuh
```

```

> MMWSH = (StrS * (StrQ/StrZ))/StrH #Nilai MMWS Strata 2
yang Tidak Patuh
> MMWSI = (StrT * (StrQ/StrZ))/StrI #Nilai MMWS Strata 3
yang Tidak Patuh

#Tabel MMWS
> MMWS = data.frame(Strata=c('1','2','3'),
  Patuh=c(MMWSA,MMWSB,MMWSC),
  TidakPatuh=c(MMWSG,MMWSH,MMWSI))

```

Hasil perhitungan nilai MMWS dan pembentukan strata baru dapat dilihat pada tabel 5.7 dan 5.8.

Tabel 5.7 Nilai MMWS disetiap Strata

Strara	Sub-Strata	MMWS	
		Patuh	Tidak Patuh
3	1	2,238	0,623
	2	1,010	0,989
	3	0,647	2,469

Tabel 5.8 Nilai setiap strata setelah dilakukan pembobotan

Strata	Sub-Strata	Patuh	Tidak Patuh	Jumlah responden
3	1	16	14	30
	2	15	14	29
	3	16	15	31

Tabel 5.8 memberikan informasi bahwa setelah dilakukan pembobotan, responden menyebar rata pada kelompok treatment (patuh) dan kontrol (tidak patuh). Berbeda dengan strata sebelum pembobotan (lihat Tabel 5.6).

8. Melakukan pemeriksaan *balancing* terhadap kovariat disetiap strata. Pengujian balance menggunakan uji t dua sampel dengan hipotesis penelitian adalah sebagai berikut:

H_0 : Tidak terdapat perbedaan mean antara kelompok patuh dan kelompok tidak patuh

H_1 : Terdapat perbedaan mean antara kelompok kelompok patuh dan kelompok tidak patuh

Daerah kritis : tolak H_0 jika $t_{stat} > t_{tabel}$ atau $p\text{-value} < \alpha$. Dengan taraf signifikansi $\alpha = 10\%$.

Berikut ini perintah yang digunakan di R untuk menguji balance kovariat.

```
#Uji T dan Balancing
##Psil
> T11 = if (var(PropensityDF$Psil[PropensityDF$Strata
=="1" & PropensityDF$Kepatuhan == "1"])==0 &
var(PropensityDF$Psil[PropensityDF$Strata == "1" &
PropensityDF$Kepatuhan == "2"])==0) {"NOL"} else
{t.test(PropensityDF$Psil[PropensityDF$Strata == "1"
& PropensityDF$Kepatuhan == "1"],
PropensityDF$Psil[PropensityDF$Strata == "1" &
PropensityDF$Kepatuhan == "2"])}

> T12 = if (var(PropensityDF$Psil[PropensityDF$Strata
=="2" & PropensityDF$Kepatuhan == "1"])==0 &
var(PropensityDF$Psil[PropensityDF$Strata == "2" &
PropensityDF$Kepatuhan == "2"])==0) {"NOL"}
else {t.test(PropensityDF$Psil[PropensityDF$Strata
=="2" & PropensityDF$Kepatuhan == "1"],
PropensityDF$Psil[PropensityDF$Strata == "2" &
PropensityDF$Kepatuhan == "2"])}

> T13 = if (var(PropensityDF$Psil[PropensityDF$Strata
=="3" & PropensityDF$Kepatuhan == "1"])==0 &
var(PropensityDF$Psil[PropensityDF$Strata == "3" &
PropensityDF$Kepatuhan == "2"])==0) {"NOL"}
else {t.test(PropensityDF$Psil[PropensityDF$Strata
=="3" & PropensityDF$Kepatuhan == "1"],
PropensityDF$Psil[PropensityDF$Strata == "3" &
PropensityDF$Kepatuhan == "2"])}

```



```

##Psi2
> T21 = if (var(PropensityDF$Psi2[PropensityDF$Strata
=="1" & PropensityDF$Kepatuhan == "1"])==0 &
var(PropensityDF$Psi2[PropensityDF$Strata == "1" &
PropensityDF$Kepatuhan == "2"])==0) {"NOL"} else
{t.test(PropensityDF$Psi2[PropensityDF$Strata == "1"
& PropensityDF$Kepatuhan == "1"],
PropensityDF$Psi2[PropensityDF$Strata == "1" &
PropensityDF$Kepatuhan == "2"])}

> T22 = if (var(PropensityDF$Psi2[PropensityDF$Strata
=="2" & PropensityDF$Kepatuhan == "1"])==0 &
var(PropensityDF$Psi2[PropensityDF$Strata == "2" &
PropensityDF$Kepatuhan == "2"])==0) {"NOL"} else
{t.test(PropensityDF$Psi2[PropensityDF$Strata == "2"
& PropensityDF$Kepatuhan == "1"],
PropensityDF$Psi2[PropensityDF$Strata == "2" &
PropensityDF$Kepatuhan == "2"])}

> T23 = if (var(PropensityDF$Psi2[PropensityDF$Strata
=="3" & PropensityDF$Kepatuhan == "1"])==0 &
var(PropensityDF$Psi2[PropensityDF$Strata == "3" &
PropensityDF$Kepatuhan == "2"])==0) {"NOL"} else
{t.test(PropensityDF$Psi2[PropensityDF$Strata == "3"
& PropensityDF$Kepatuhan == "1"],
PropensityDF$Psi2[PropensityDF$Strata == "3" &
PropensityDF$Kepatuhan == "2"])}

> BalanPsi1 = cbind(round(T11$p.value,3),
round(T12$p.value,3),round(T13$p.value,3))

> BalanPsi2 = cbind(round(T21$p.value,3),
round(T22$p.value,3),round(T23$p.value,3))

```

Hasil pengujian balance yang diperoleh adalah sebagai berikut.

Tabel 5.9 Uji keseimbangan kovariat

Strata	Sub-Strata	Kovariat 1		Kovariat 2	
		t-stat	Keputusan	t-stat	Keputusan
3	1	1,5179	Gagal Tolak H_0	1,1916	Gagal Tolak H_0
	2	0,2012	Gagal Tolak H_0	0,3263	Gagal Tolak H_0
	3	1,8846	Gagal Tolak H_0	0,4615	Gagal Tolak H_0

Tabel 5.9 menunjukkan bahwa kovariat 1 dan 2 sudah balance untuk semua strata. Hal ini dapat dilihat dari keputusan gagal tolak H_0 karena $t\text{-stat} (t\text{-hitung}) < t_{\alpha/2, n-2} = 1,96$ untuk $\alpha = 5\%$. Balance yang dimaksudkan pada tahapan ini adalah tidak terdapat perbedaan mean antara kelompok treatment dan kontrol.

9. Menghitung *effect-treatment* dengan menggunakan metode MMWS

Berikut ini syntax R yang digunakan untuk estimasi ATE berdasarkan persamaan (1.5) sampai (1.7).

```
#ATT dan ATE PS2
> EAggS_PS2 = aggregate(PropensityDF_PS2$e,
  list(PropensityDF_PS2$Strata_2,
  PropensityDF_PS2$Kepatuhan), sum, simplify = TRUE)
#jumlah e untuk tiap subgroup

> Y3z1_PS2 = EAggS_PS2[1, 3] #Y1 (z)
> Y2z1_PS2 = EAggS_PS2[2, 3] #Y2 (z)
> Y1z1_PS2 = EAggS_PS2[3, 3] #Y3 (z)

> Y1z2_PS2 = sum(PropensityDF_PS2$e[which
  (PropensityDF_PS2$Strata_2 ==1 &
  PropensityDF_PS2$Kepatuhan==2)] * (StrA_PS2/StrG_PS2))

:
:
```

```

> Y3z2_PS2=sum(PropensityDF_PS2$e[which
  PropensityDF_PS2$Strata_2==3 &
  PropensityDF_PS2$Kepatuhan==2])* (StrC_PS2/StrI_PS2))

#ATE
> MEt_PS2 = mean(PropensityDF_PS2$e[which
  (PropensityDF_PS2$Kepatuhan==1)])
> VEt_PS2 = var(PropensityDF_PS2$e[which
  (PropensityDF_PS2$Kepatuhan==1)])
> MEc_PS2 = mean(PropensityDF_PS2$e[which
  (PropensityDF_PS2$Kepatuhan==2)])
> Vec_PS2 = var(PropensityDF_PS2$e[which
  (PropensityDF_PS2$Kepatuhan==2)])

> ATT_t_PS2 = (Y1z1_PS2+Y2z1_PS2+Y3z1_PS2)/length(Nmr)
> ATT_c_PS2 = (Y1z2_PS2+Y2z2_PS2+Y3z2_PS2)/length(Nmr)
  #Jika ATT atau ATE tdk sesuai manual,cek Yiz1 dan Yiz2
> ATE_PS2 = abs (ATT_t_PS2 - ATT_c_PS2)
> SE_PS2      = sqrt(( var(PropensityDF_PS2$e
  [PropensityDF_PS2$Kepatuhan == "1"]) /
  length(PropensityDF_PS2$e
  [PropensityDF_PS2$Kepatuhan == "1"])) +
  var(PropensityDF_PS2$e
  [PropensityDF_PS2$Kepatuhan == "2"]) /
  length(PropensityDF_PS2$e
  [PropensityDF_PS2$Kepatuhan == "2"]))

> TETot_PS2 = if (var(PropensityDF_PS2$e
  [PropensityDF_PS2$Kepatuhan == "1"])==0 &
  var(PropensityDF_PS2$e
  [PropensityDF_PS2$Kepatuhan == "2"])==0) {"NOL"} else
  {t.test(PropensityDF_PS2$e
  [PropensityDF_PS2$Kepatuhan == "1"],
  PropensityDF_PS2$e[PropensityDF_PS2$Kepatuhan== "2"])}

> RESPONDEN = data.frame(Responden=c(Nmr),PS=c(e))
> STRATA = data.frame(Strata=c('1','2','3','Total'),
  Treatment=c(StrA,StrB,StrC,StrP),
  Control=c(StrG,StrH,StrI,StrQ),

```

```

> Jumlah_Responden=c(StrR,StrS,StrT,StrZ)
> MMWS = data.frame(Strata=c('1','2','3'),
                    Treatment=c(MMWSA,MMWSB,MMWSC),
                    Control=c(MMWSG,MMWSH,MMWSI))
> NEW_RESPONDEN = data.frame(Responden=
  c(PropensityDF_PS2$Nmr),
  PS_2=c(PropensityDF_PS2$e))
> NEW_STRATA = data.frame(Strata=c('1','2','3','Total'),
  Treatment=c(StrA_PS2,StrB_PS2,StrC_PS2,StrP_PS2),
  Control=c(StrG_PS2,StrH_PS2,StrI_PS2,StrQ_PS2),
  Jumlah_Responden=c(StrR_PS2,StrS_PS2,StrT_PS2,
  StrZ_PS2))
> BALENCING = data.frame(Strata=c('1','2','3'),
  p_value_Kov1=c(T11$p.value,T12$p.value,T13$p.value),
  p_value_Kov2=c(T21$p.value,T22$p.value,T23$p.value))
> ATE = data.frame(ATE=c(ATE_PS2),SE=c(SE_PS2),
  p_value=c(TETot_PS2$p.value))

```

Hasil estimasi ATE disajikan pada Tabel 5.10.

Tabel 5.10 Tabel Nilai ATE dari *propensity score*

Strata	ATE	SE	<i>t-stat</i>	<i>t-tabel</i>
3	0,0275	0,0376	1,3174	1,2912

Hipotesis dalam bidang kesehatan umumnya menggunakan taraf signifikansi $\alpha = 5\%$. Akan tetapi pada kasus ini jika menggunakan $\alpha = 5\%$ maka *effect treatment* di 3 strata tidak signifikan. Oleh karena penggunaan $\alpha = 5\%$ menyebabkan *effect treatment* tidak signifikan, maka pada penelitian ini menggunakan taraf signifikansi $\alpha = 20\%$. Untuk 3 strata diperoleh *effect treatment* sebesar 0,0275 dengan nilai $t_{\text{-stat}}$ (1,3177) > $t_{\text{-tabel}}$ (1,2912) sehingga tolak H_0 . Ini berarti bahwa faktor *confounding* kepatuhan terapi ARV berpengaruh signifikan terhadap terjadinya infeksi oportunistik pada pasien HIV/AIDS.

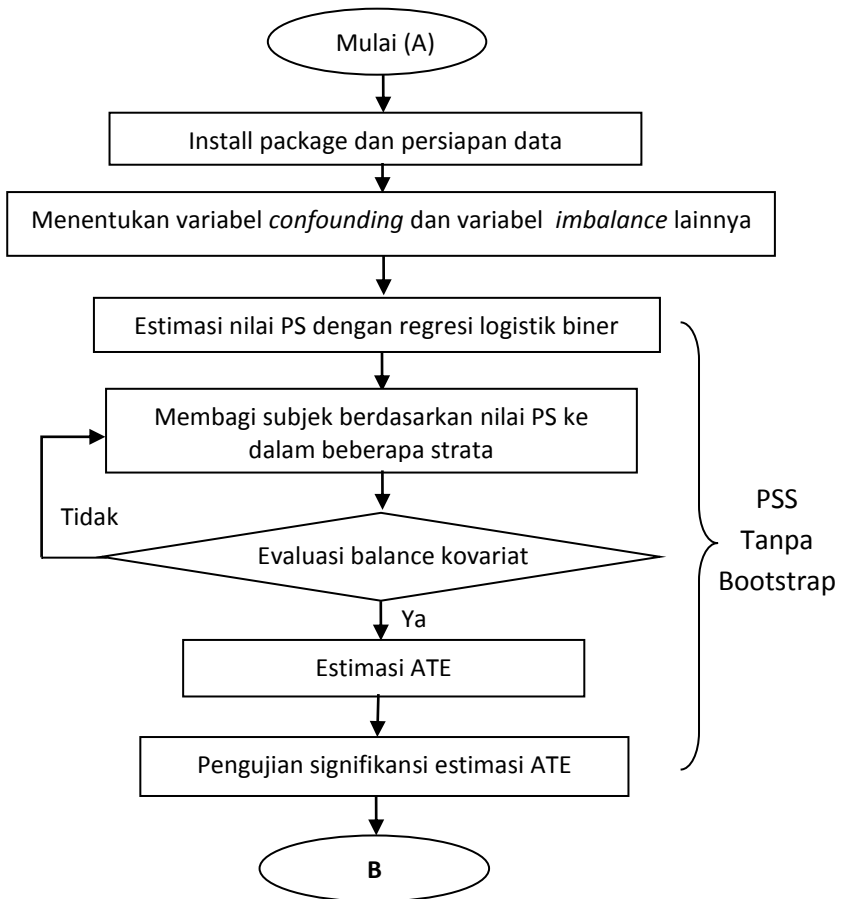
Propensity Score Stratification Bootstrap Using R Language

Pada dasarnya PSSB merupakan lanjutan dari PSS tanpa bootstrap. Oleh karena itu, akan digunakan kasus yang sama dengan PSS yang telah dibahas pada pembahasan sebelumnya.

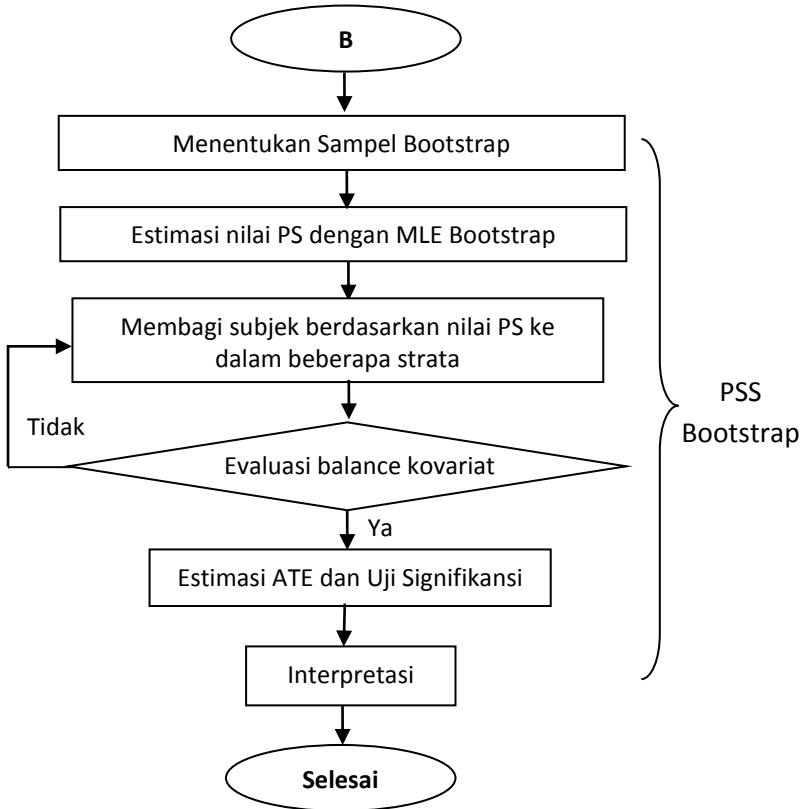
Diberikan Data pasien DM untuk kasus Neuropati Diabetik Perifer (NDP) yang terdiri atas 7 variabel yakni 1 variabel respon berupa Status NDP pasien DM, dan 6 variabel bebas (kovariat) berupa Umur pasien (X1), lama menderita DM (X2), jenis kelamin (X3), Obesitas (X4), Hipertensi (X5), Kadar Asam Urat Serum (X6). Peneliti ingin menerapkan metode propensity score stratification bootstrap pada data ini. Mengingat banyak langkah yang dilakukan untuk sampai pada tahapan estimasi efek treatment, maka strata yang dibuat pada kasus ini hanya 2 strata saja. Untuk mengetahui lebih detail terkait dengan penerapan PSSB pada data NDP, pada pembahasan ini akan diuraikan cara pengerjaan manualnya menggunakan software R.

Pada Pembahasan ini akan digunakan informasi yang diperoleh pada PSS untuk kasus yang sama yang telah

dibahas sebelumnya. Pada pembahasan metode PSS tanpa bootstrap diketahui bahwa kovariat yang dijadikan sebagai variabel *confounding* adalah kadar Asam Urat Serum (AUS). Variabel kadar AUS ini akan dijadikan juga sebagai *confounding* pada metode ini. Setelah variabel *confounding* ditentukan, maka langkah selanjutnya adalah estimasi nilai propensity score. Untuk lebih jelas, tahapan PSSB pada kasus ini dapat dilihat pada Gambar 6.1.



Gambar 6.1 Tahapan PSS Bootstrap untuk kasus DM



Gambar 6.1 Tahapan PSS Bootstrap untuk kasus DM (lanjutan)

- A. Analisis Data menggunakan Propensity Score Stratification (PSS) tanpa Bootstrap
1. Install package yang digunakan

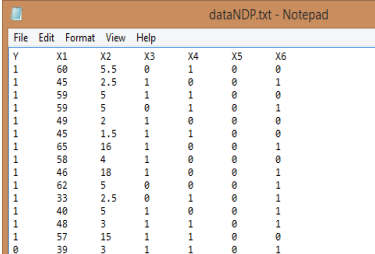
Untuk menggunakan package boot, terlebih dahulu install package dari menu package, jika sudah diinstall selanjutnya load package dengan perintah berikut.

```
## memanggil package yang digunakan
library(boot)
```

2. Mempersiapkan data yang digunakan

Untuk memulai tahapan propensity score stratification diperlukan data terkait kasus yang diteliti. Data disimpan dalam notepad (.txt) dengan format variabel sebagai kolom dan observasi (subjek) sebagai baris.

Berikut contoh format data dalam notepad



Y	X1	X2	X3	X4	X5	X6
1	68	5.5	0	1	0	0
1	45	2.5	1	0	0	1
1	59	5	1	1	0	0
1	59	5	0	1	0	1
1	49	2	1	0	0	0
1	45	1.5	1	1	0	0
1	65	16	1	0	0	1
1	58	4	1	0	0	0
1	46	18	1	0	0	1
1	62	5	0	0	0	1
1	33	2.5	0	1	0	1
1	40	5	1	0	0	1
1	48	3	1	1	0	1
1	57	15	1	1	0	0
0	39	3	1	1	0	1

Selanjutnya mendefinisikan kovariat yang bertipe kategori (kecuali X6 yang ditetapkan sebagai variabel *confounding*) sebagai faktor.

```
#read data
> data_NDP<-read.table("D:/lain-lain/dataNDP.txt",
  header = TRUE)
> data_NDP$X3<-factor(data_NDP$X3)
> data_NDP$X4<-factor(data_NDP$X4)
> data_NDP$X5<-factor(data_NDP$X5)
```


3. Memilih kovariat sebagai variabel *confounding* dan kovariat *imbalance*

Pemilihan variabel *confounding* didasarkan pada teori dan bukti empiris tentang hubungan antara kovariat. Berdasarkan pembahasan pada bab-4 tentang PSS pada data yang sama (DM), diketahui bahwa kovariat yang dijadikan variabel *confounding* adalah kadar asam urat serum (X6) dan kovariat yang *imbalance* adalah X2 (lama DM).

4. Estimasi Nilai Propensity Score dengan Metode MLE

Estimasi nilai propensity yang diperoleh pada sub-bab ini hampir sama dengan estimasi nilai propensity yang dibahas pada bab-4, akan tetapi nilai propensity score yang dihasilkan sedikit berbeda, namun tidak merubah hasil pada tahapan stratifikasi sampai estimasi ATE. Berikut ini estimasi nilai propensity menggunakan software R.

```
> attach(data_NDP)
> ps_logit <- glm(X6~X1+X2+X3+X4+X5,data=data_NDP,
  family=binomial("logit"))
> summary(ps_logit)
> data_NDP$ps = ps_logit$fitted #nilai propensity

# analisis data asli
> coefs <- ps_logit$coef #parameter beta
> se <- sqrt(diag(vcov(ps_logit))) #parameter se
> z <- coefs/se
```

Statistik uji z pada perintah diatas diperoleh berdasarkan persamaan (1.19).

5. Stratifikasi dan Pengujian Balance Kovariat

Stratifikasi dan pengujian balance kovariat sama dengan pembahasan sebelumnya (lihat bab-4), akan tetapi pada bab ini subjek dibagi menjadi dua strata saja. Pada kasus ini, balance kovariat diuji menggunakan uji z untuk data kategori dan uji t untuk data kontinu. Berikut ini perintah yang digunakan di R untuk stratifikasi dan pengujian balance kovariat.

```
### uji balance sebelum strata
## untuk X1 dan X2
> tx1=t.test(data_NDP$X1[data_NDP$X6 == "0"],
  data_NDP$X1[data_NDP$X6 == "1"])
> tx2=t.test(data_NDP$X2[data_NDP$X6 == "0"],
  data_NDP$X2[data_NDP$X6 == "1"])
#untuk X3
> ax3_1=data_NDP$X3[(data_NDP$X3=="0" &
  data_NDP$X6 == "0")]
> ax3_2=data_NDP$X3[(data_NDP$X3=="1" &
  data_NDP$X6 == "0")]
> ax3_3=data_NDP$X3[(data_NDP$X3=="0" &
  data_NDP$X6 == "1")]
> ax3_4=data_NDP$X3[(data_NDP$X3=="1" &
  data_NDP$X6 == "1")]
> X3tes = as.table(rbind(c(length((ax3_1)),
  length((ax3_2))),c(length((ax3_3)),length((ax3_4)))))
> x3_0=length((ax3_1))
> n3_0=length((ax3_1))+length((ax3_2))
> x3_1=length((ax3_3))
> n3_1=length((ax3_3))+length((ax3_4))
> z.propx3 = function(x3_0,x3_1,n3_0,n3_1){
  numeratorx3 = ((x3_0)/(n3_0)) - ((x3_1)/(n3_1))
  p.commonx3 = ((x3_0)+(x3_1)) / ((n3_0)+(n3_1))
  denominatorx3 = sqrt(p.commonx3 * (1-p.commonx3) *
  (1/(n3_0) + 1/(n3_1)))
  z.prop.risx3 = numeratorx3/denominatorx3
  return(z.prop.risx3)}
> zx3=z.propx3(x3_0,x3_1,n3_0,n3_1) #memanggil z score
> p.value_x3=2*pnorm(-abs(zx3))
```

```

#untuk X4
> ax4_1=data_NDP$X4[(data_NDP$X4=="0" &
  data_NDP$X6 == "0")]
> ax4_2=data_NDP$X4[(data_NDP$X4=="1" &
  data_NDP$X6 == "0")]
> ax4_3=data_NDP$X4[(data_NDP$X4=="0" &
  data_NDP$X6 == "1")]
> ax4_4=data_NDP$X4[(data_NDP$X4=="1" &
  data_NDP$X6 == "1")]
> X4tes = as.table(rbind(c(length((ax4_1)),
  length((ax4_2))), c(length((ax4_3)),length((ax4_4))))
> colnames(X4tes) <- rownames(X4tes) <- c("0", "1")
> names(dimnames(X4tes)) = c("X8", "X4")
> x4_0=length((ax4_1))
> n4_0=length((ax4_1))+length((ax4_2))
> x4_1=length((ax4_3))
> n4_1=length((ax4_3))+length((ax4_4))
> z.propx4 = function(x4_0,x4_1,n4_0,n4_1){
  numeratorx4 = ((x4_0)/(n4_0)) - ((x4_1)/(n4_1))
  p.commonx4 = ((x4_0)+(x4_1)) / ((n4_0)+(n4_1))
  denominatorx4 = sqrt(p.commonx4 * (1-p.commonx4) *
  (1/(n4_0) + 1/(n4_1)))
  z.prop.risx4 = numeratorx4/denominatorx4
  return(z.prop.risx4)}
> zx4=z.propx4(x4_0,x4_1,n4_0,n4_1) #memanggil z score
> p.value_x4=2*pnorm(-abs(zx4))

#untuk X5
> ax5_1=data_NDP$X5[(data_NDP$X5=="0" &
  data_NDP$X6 == "0")]
> ax5_2=data_NDP$X5[(data_NDP$X5=="1" &
  data_NDP$X6 == "0")]
> ax5_3=data_NDP$X5[(data_NDP$X5=="0" &
  data_NDP$X6 == "1")]
> ax5_4=data_NDP$X5[(data_NDP$X5=="1" &
  data_NDP$X6 == "1")]
> X5tes = as.table(rbind(c(length((ax4_1)),

```

```

length((ax5_2))), c(length((ax5_3)),length((ax5_4))))
> colnames(X5tes) <- rownames(X5tes) <- c("0", "1")
> names(dimnames(X5tes)) = c("X6", "X5")
> x5_0=length((ax5_1))
> n5_0=length((ax5_1))+length((ax5_2))
> x5_1=length((ax5_3))
> n5_1=length((ax5_3))+length((ax5_4))
> z.propx5 = function(x5_0,x5_1,n5_0,n5_1){
  numeratorx5 = ((x5_0)/(n5_0)) - ((x5_1)/(n5_1))
  p.commonx5 = ((x5_0)+(x5_1)) / ((n5_0)+(n5_1))
  denominatorx5 = sqrt(p.commonx5 * (1-p.commonx5) *
  (1/(n5_0) + 1/(n5_1)))
  z.prop.risx5 = numeratorx5 /denominatorx5
  return(z.prop.risx5)}
> zx5=z.propx5(x5_0,x5_1,n5_0,n5_1) #memanggil z score
> p.value_x5=2*pnorm(-abs(zx5))

## test balance setelah strata untuk data kontinu __ X1
tx11=t.test(data_NDP$X1[data_NDP$strata == "1" &
data_NDP$X6 == "0"],
data_NDP$X1[data_NDP$strata == "1" & data_NDP$X6 == "1"])
tx12=t.test(data_NDP$X1[data_NDP$strata == "2" &
data_NDP$X6 == "0"],data_NDP$X1[data_NDP$strata == "2" &
data_NDP$X6 == "1"])

## test balance setelah strata untuk data kontinu __ X2
tx21=t.test(data_NDP$X2[data_NDP$strata == "1" &
data_NDP$X6 == "0"],
data_NDP$X2[data_NDP$strata == "1" & data_NDP$X6 == "1"])
tx22=t.test(data_NDP$X2[data_NDP$strata == "2" &
data_NDP$X6 == "0"],data_NDP$X2[data_NDP$strata == "2" &
data_NDP$X6 == "1"])

:
## test balance setelah strata untuk data kategori __ X5
#stratal
> ax51_1=data_NDP$X5[(data_NDP$X5=="0" &
data_NDP$strata == "1" & data_NDP$X6 == "0")]
> ax51_2=data_NDP$X5[(data_NDP$X5=="1" &
data_NDP$strata == "1" & data_NDP$X6 == "0")]
> ax51_3=data_NDP$X5[(data_NDP$X5=="0" &

```

```

data_NDP$strata == "1" & data_NDP$X6 == "1")
> ax51_4=data_NDP$X5[(data_NDP$X5=="1" &
data_NDP$strata == "1" & data_NDP$X6 == "1")]
> X5tes1 = as.table(rbind(c(length((ax51_1)),
length((ax51_2))),c(length((ax51_3)),
length((ax51_4))))
> colnames(X5tes1) <- rownames(X5tes1) <- c("0", "1")
> names(dimnames(X5tes1)) = c("X6", "X5")
> x5_01=length((ax51_1))
> n5_01=length((ax51_1))+length((ax51_2))
> x5_11=length((ax51_3))
> n5_11=length((ax51_3))+length((ax51_4))
> z.propx51 = function(x5_01,x5_11,n5_01,n5_11){
numeratorx51 = ((x5_01)/(n5_01)) - ((x5_11)/(n5_11))
p.commonx51 = ((x5_01)+(x5_11)) / ((n5_01)+(n5_11))
denominatorx51 = sqrt(p.commonx51 * (1-p.commonx51) *
(1/(n5_01) + 1/(n5_11)))
z.prop.risx51 = numeratorx51 /denominatorx51
return(z.prop.risx51)}
> zx51=z.propx51(x5_01,x5_11,n5_01,n5_11) #call z score

#strata2
> ax52_1=data_NDP$X5[(data_NDP$X5=="0" &
data_NDP$strata == "2" & data_NDP$X6 == "0")]
> ax52_2=data_NDP$X5[(data_NDP$X5=="1" &
data_NDP$strata == "2" & data_NDP$X6 == "0")]
> ax52_3=data_NDP$X5[(data_NDP$X5=="0" &
data_NDP$strata == "2" & data_NDP$X6 == "1")]
> ax52_4=data_NDP$X5[(data_NDP$X5=="1" &
data_NDP$strata == "2" & data_NDP$X6 == "1")]
> X5tes2 = as.table(rbind(c(length((ax52_1)),
length((ax52_2))),c(length((ax52_3)),
length((ax52_4))))
> colnames(X5tes2) <- rownames(X5tes2) <- c("0", "1")
> names(dimnames(X5tes2)) = c("X6", "X5")
> x5_02=length((ax52_1))
> n5_02=length((ax52_1))+length((ax52_2))
> x5_12=length((ax52_3))

```

```

> n5_12=length((ax52_3))+length((ax52_4))
> z.propx52 = function(x5_02,x5_12,n5_02,n5_12){
  Numeratorx52 = ((x5_02)/(n5_02)) - ((x5_12)/(n5_12))
  p.commonx52 = ((x5_02)+(x5_12)) / ((n5_02)+(n5_12))
  denominatorx52 = sqrt(p.commonx52 * (1-p.commonx52) *
    (1/(n5_02) + 1/(n5_12)))
  z.prop.risx52 = numeratorx52 /denominatorx52
  return(z.prop.risx52)}
> zx52=z.propx52(x5_02,x5_12,n5_02,n5_12)#fungsi z score

#summary balance
> summary.bal.strat2=cbind(c(tx1$p.value,tx11$p.value,
  tx12$p.value),c(tx2$p.value,tx21$p.value,
  tx22$p.value),c(p.value_x3,p.value_x31,p.value_x32),
  c(p.value_x4,p.value_x41,p.value_x42),
  c(p.value_x5,p.value_x51,p.value_x52))
> colnames(summary.bal.strat2)=c("X1","X2","X3",
  "X4","X5")
> rownames(summary.bal.strat2) <- c("before strata",
  "strata 1", "strata 2")
> summary.bal.strat2

```

dan diperoleh hasil pengujian balance sebagai berikut.

	X1	X2	X3
before strata	0.5547476	0.004888416	0.03076844
strata 1	0.8823585	0.347517015	0.87882781
strata 2	0.4572331	0.534169440	0.27967256
	X4	X5	
before strata	0.5369626	0.0005003408	
strata 1	0.7351517	0.1252456588	
strata 2	0.8325063	0.5398416450	

Hasil pengujian balance menunjukkan bahwa semua kovariat sudah memenuhi asumsi balance untuk semua strata

6. Estimasi ATE dan Uji Signifikansinya

Setelah dilakukan pengujian balance, selanjutnya estimasi ATE manual. Berikut ini perintah yang digunakan untuk estimasi ATE sesuai persamaan (1.22)

```
##### ESTIMASI ATE MANUAL #####
> Y2.QT0 <- split(data_NDP$Y[data_NDP$X6==0],
  data_NDP$strata[data_NDP$X6==0])
> Y2.QT1 <- split(data_NDP$Y[data_NDP$X6==1],
  data_NDP$strata[data_NDP$X6==1])
> P=sapply(Y2.QT1, mean)      # estimates for group trt=1
> S=sapply(Y2.QT0, mean)      # estimates for group trt=0
> K=sapply(Y2.QT1, mean)-sapply(Y2.QT0, mean) #TE strata
> nc1=length(Y2.QT0[[1]])     #n control strata 1
> nc2=length(Y2.QT0[[2]])     #n control strata 2
> nt1=length(Y2.QT1[[1]])     #n treatment strata 1
> nt2=length(Y2.QT1[[2]])     #n treatment strata 2
> nc=nc1+nc2                  # n control
> nt=nt1+nt2                  # n treatment
> theta1=((nt1+nc1)*(K[1]))/(nt+nc) # ate strata 1
> theta2=((nt2+nc2)*(K[2]))/(nt+nc) # ate strata 2
> theta_hat2=(theta1+theta2)   #ATE_final

### mencari SE_ATE #####
> Svar_treat=sapply(Y2.QT1, var)      # var treatment
> Svar_control=sapply(Y2.QT0, var)    # var control
> Vt1=((1/nt1))*Svar_treat[1]         # var treat strata 1
> Vt2=((1/nt2))*Svar_treat[2]         # var treat strata 2
> Vc1=((1/nc1))*Svar_control[1]       # var control strata1
> Vc2=((1/nc2))*Svar_control[2]       # var control strata2
> var1=Vt1+Vc1                       # var strata 1
> var2=Vt2+Vc2                       # var strata 2
> Var_theta1=((nt1+nc1)^2)*(var1)     # var_ate strata 1
> Var_theta2=((nt2+nc2)^2)*(var2)     # var_ate strata 2
> Var_theta_hat=(Var_theta1+Var_theta2)/((nt+nc)^2)
> se_theta_hat2=sqrt(Var_theta_hat)   # SE ATE
```

```

## uji t stat ##
> t2=(theta_hat2)/(se_theta_hat2)

## calculating estimate of df ##
> numstrata=2
> df2 = length((Y)) - (2*numstrata)

## calculating p-value for t test ##
> p_value2=pt(-abs(t2),df2)

## print output
> theta_hat2
> se_theta_hat2
> t2
> p_value2

```

Berdasarkan syntax diatas, diperoleh estimasi ATE dan hasil uji signifikansi menggunakan uji t seperti ditunjukkan pada Tabel berikut.

```

> theta_hat2
      1
0.4738336

> se_theta_hat2
      1
0.1075199

> t2
      1
4.406937

> p_value2
      1
1.652954e-05

```


Output diatas menunjukkan bahwa estimasi ATE adalah sebesar 0,4738 dengan standar error sebesar 0,1075. Pada uji signifikansi ATE diperoleh nilai $p\text{-value} < \alpha = 5\%$. Hal ini menunjukkan bahwa kadar asam urat serum tinggi berpengaruh signifikan terhadap kejadian NDP dengan pengaruh sebesar 0,4738.

B. Analisis Data menggunakan Propensity Score Stratification (PSS) Bootstrap

1. Estimasi Nilai Propensity Score dengan Metode MLE Bootstrap

Metode bootstrap pada software R menggunakan package khusus yang membahas tentang bootstrap yakni package `boot`. Pada metode ini salah satu hal penting adalah banyak replikasi yang dilakukan untuk estimasi. Untuk kasus DM digunakan 200 replikasi. Estimasi parameter bootstrap menggunakan persamaan (1.17).

Sebelum estimasi dengan metode bootstrap, terlebih dahulu dilakukan estimasi parameter pada full data, hal ini disebabkan karena kriteria penolakan H_0 menggunakan nilai $p\text{-value}$ (persamaan 1.22) yang melibatkan statistik uji pada full data. Berikut ini perintah estimasi parameter untuk full data menggunakan package **glm**.

Setelah diperoleh statistik uji untuk full data (data tanpa bootstrap), selanjutnya dilakukan estimasi parameter menggunakan package **boot** dengan perintah sebagai berikut.

```
#bootstrap dengan package  
> library(boot)
```

```

> logit.bootstrap <- function(data_NDP, indices) {
  d <- data_NDP[indices, ]
  fit <- glm(X6~X1+X2+X3+X4+X5, data = d,
    family = "binomial")
  return(coef(fit))}
> set.seed(2000)
> logit.boot <- boot(data=data_NDP,
  statistic=logit.bootstrap,R=200)
> b0=mean(logit.boot$t[,1]) #parameter intercept
> b1=mean(logit.boot$t[,2]) #parameter x1
:
:
> b5=mean(logit.boot$t[,6]) #parameter x5
> se_boot=sqrt(diag(cov(logit.boot$t)))

```

Nilai b_0 sampai b_5 masing-masing merupakan estimasi parameter intercept sampai parameter X_5 . Nilai estimasi diperoleh menggunakan persamaan (1.30). setelah diperoleh nilai estimasi parameter bootstrap, Selanjutnya dilakukan pengujian signifikansi parameter menggunakan uji Z^* bootstrap sesuai persamaan (1.20) dan menghitung nilai *p-value* sesuai persamaan (1.24). Perintah yang digunakan untuk menguji signifikansi parameter pada software R adalah sebagai berikut.

```

## p-value
> z1.under.h0=(logit.boot$t[,1]-
  logit.boot$t0[1])/se_boot[1]
> z2.under.h0=(logit.boot$t[,2]-
  logit.boot$t0[2])/se_boot[2]
:
:
> z6.under.h0=(logit.boot$t[,6]-
  logit.boot$t0[8])/se_boot[6]

> p.valu1 = 1/(1+logit.boot$R) *
  (1+sum(ifelse(abs(z1.under.h0) >= abs(z[1]),1,0)))
> p.valu2 = 1/(1+logit.boot$R) *
  (1+sum(ifelse(abs(z2.under.h0) >= abs(z[2]),1,0)))
:
:

```

```

> p.valu6 = 1/(1+logit.boot$R) *
(1+sum(ifelse(abs(z6.under.h0) >= abs(z[6]),1,0)))

> summary.logbootx_z=cbind(c(b0,b1,b2,b3,b4,b5),
c(se_boot[1],se_boot[2],se_boot[3],
se_boot[4], se_boot[5], se_boot[6]), c(p.valu1,
p.valu2, p.valu3,p.valu4, p.valu5,p.valu6))
> colnames(summary.logbootx_z) <-
c("parameter boot", "std.error","p_value.boot")
> rownames(summary.logbootx_z) <- c("intercept",
"X1","X2","X31","X41","X51")

```

Berdasarkan perintah diatas, diperoleh hasil pengujian parameter bootstrap sebagai berikut.

```

> summary.logbootx_z

```

	parameter boot	std.error	p_value.boot
intercept	-0.02129353	2.41311526	0.96019900
X1	-0.02193475	0.04437778	0.58208955
X2	0.15106950	0.09572770	0.03980100
X31	-0.80782168	0.72673725	0.15422886
X41	-0.02245063	0.56367131	0.86567164
X51	1.34624212	0.71989665	0.05472637

Output diatas menunjukkan nilai propensity score sebagai berikut.

$$e(X_i)^* = \frac{\exp(b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5)}{1 + \exp(b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5)} \quad (1.1)$$

Selain itu, uji signifikansi parameter bootstrap memberikan gambaran bahwa pada taraf signifikansi $\alpha = 10\%$ terdapat dua kovariat yang signifikan, yakni kovariat pasien DM menderita hipertensi dan lama pasien menderita DM.

C. Stratifikasi, Mengecek Balance Kovariat dan Estimasi ATE

Stratifikasi, cek balance kovariat dan estimasi ATE pada sampel bootstrap menggunakan package **boot**. Akan tetapi sebelum melakukan stratifikasi, terlebih dahulu kovariat yang didefinisikan sebagai faktor pada full data, diubah menjadi integer dengan memanfaatkan package **varhandle** dengan perintah **unfactor** sesuai yang ditunjukkan pada tabel berikut. Setelah itu dihitung estimasi nilai propensity score berdasarkan estimasi parameter bootstrap yang telah diperoleh pada tahapan sebelumnya. Selanjutnya mendefinisikan fungsi statistic yang akan dibootstrapping. Fungsi statistic berisi perhitungan manual uji balance kovariat masing-masing strata dan estimasi ATE manual. Untuk lebih jelasnya dapat dilihat pada tabel berikut ini.

```
> require(varhandle)
> data_NDP$X3<- unfactor(data_NDP$X3)
> data_NDP$X4<- unfactor(data_NDP$X4)
> data_NDP$X5<- unfactor(data_NDP$X5)
> numlog<-exp(b0+b1*data_NDP$X1+b2*data_NDP$X2+
  b3*data_NDP$X3+b4* data_NDP$X4+b5*data_NDP$X5)
> data_NDP$ps1=(numlog/(1+numlog)) #nilai propensity
> fn.boot.bedamean2 <- function(data_NDP, indices){
  newdata <- data_NDP[indices,]
  breakvals = quantile(newdata$ps1,
  prob = seq(from = 0,to = 1, by = 1/2), na.rm = T)
  newdata$strat1 = cut(newdata$ps1,
  breaks=breakvals,labels=c('1','2'),
  include.lowest=TRUE)
### uji balance sebelum strata
## untuk X1
x1_0= newdata$X1[newdata$X8 == "0"]
x1_1= newdata$X1[newdata$X8 == "1"]
n1_0=length(newdata$X1[newdata$X8 == "0"])
n1_1=length(newdata$X1[newdata$X8 == "1"])
tttest <- function(m1,m2,s1,s2,n1,n2,m0=0,
```

```

equal.variance=FALSE)
{diff_mean <- (m1-m2-m0)
  return(diff_mean)}
tx1=ttest( mean(x1_0), mean(x1_1), sd(x1_0), sd(x1_1),
n1_0, n1_1)

#untuk X2
x2_0= newdata$X2[newdata$X6 == "0"]
x2_1= newdata$X2[newdata$X6 == "1"]
n2_0=length(newdata$X2[newdata$X6 == "0"]) #nc strata
n2_1=length(newdata$X2[newdata$X6 == "1"]) #nt strata
ttest <- function(m1,m2,s1,s2,n1,n2,m0=0,
equal.variance=FALSE)
{diff_mean <- (m1-m2-m0)
  return(diff_mean)}
tx2=ttest( mean(x2_0), mean(x2_1), sd(x2_0), sd(x2_1),
n2_0, n2_1)

#untuk X3
ax3_1=newdata$X3[(newdata$X3=="0" &
newdata$X6 == "0")]
ax3_2=newdata$X3[(newdata$X3=="1" &
newdata$X6 == "0")]
ax3_3=newdata$X3[(newdata$X3=="0" &
newdata$X6 == "1")]
ax3_4=newdata$X3[(newdata$X3=="1" &
newdata$X6 == "1")]
X3tes = as.table(rbind(c(length((ax3_1)),
length((ax3_2))),c(length((ax3_3)),length((ax3_4)))))
x3_0=length((ax3_1))
n3_0=length((ax3_1))+length((ax3_2))
x3_1=length((ax3_3))
n3_1=length((ax3_3))+length((ax3_4))
z.propx3 = function(x3_0,x3_1,n3_0,n3_1){
numeratorx3 = ((x3_0)/(n3_0)) - ((x3_1)/(n3_1))
p.commonx3 = ((x3_0)+(x3_1)) / ((n3_0)+(n3_1))
denominatorx3 = sqrt(p.commonx3 * (1-p.commonx3) *
(1/(n3_0) + 1/(n3_1)))

```

```

z.prop.risx3 = numeratorx3
return(z.prop.risx3)
zx3=z.propx3(x3_0,x3_1,n3_0,n3_1) #memanggil z score
p.value_x3=2*pnorm(-abs(zx3))

#untuk X4
ax4_1=newdata$X4[(newdata$X4=="0" &
newdata$X6 == "0")]
ax4_2=newdata$X4[(newdata$X4=="1" &
newdata$X6 == "0")]
ax4_3=newdata$X4[(newdata$X4=="0" &
newdata$X6 == "1")]
ax4_4=newdata$X4[(newdata$X4=="1" &
newdata$X6 == "1")]
X4tes = as.table(rbind(c(length((ax4_1)),
length((ax4_2))), c(length((ax4_3)),length((ax4_4))))
colnames(X4tes) <- rownames(X4tes) <- c("0", "1")
names(dimnames(X4tes)) = c("X8", "X4")
x4_0=length((ax4_1))
n4_0=length((ax4_1))+length((ax4_2))
x4_1=length((ax4_3))
n4_1=length((ax4_3))+length((ax4_4))
z.propx4 = function(x4_0,x4_1,n4_0,n4_1){
numeratorx4 = ((x4_0)/(n4_0)) - ((x4_1)/(n4_1))
p.commonx4 = ((x4_0)+(x4_1)) / ((n4_0)+(n4_1))
denominatorx4 = sqrt(p.commonx4 * (1-p.commonx4) *
(1/(n4_0) + 1/(n4_1)))
z.prop.risx4 = numeratorx4
return(z.prop.risx4)
}
zx4=z.propx4(x4_0,x4_1,n4_0,n4_1) #memanggil z score
p.value_x4=2*pnorm(-abs(zx4))

#untuk X5
ax5_1=newdata$X5[(newdata$X5=="0" &
newdata$X6 == "0")]
ax5_2=newdata$X5[(newdata$X5=="1" &
newdata$X6 == "0")]
ax5_3=newdata$X5[(newdata$X5=="0" &
newdata$X6 == "1")]

```

```

ax5_4=newdata$X5[(newdata$X5=="1" &
newdata$X6 == "1")]
X5tes = as.table(rbind(c(length((ax4_1)),
length((ax5_2))), c(length((ax5_3)),length((ax5_4))))
colnames(X5tes) <- rownames(X5tes) <- c("0", "1")
names(dimnames(X5tes)) = c("X6", "X5")
x5_0=length((ax5_1))
n5_0=length((ax5_1))+length((ax5_2))
x5_1=length((ax5_3))
n5_1=length((ax5_3))+length((ax5_4))
z.propx5 = function(x5_0,x5_1,n5_0,n5_1){
Numeratorx5 = ((x5_0)/(n5_0)) - ((x5_1)/(n5_1))
p.commonx5 = ((x5_0)+(x5_1)) / ((n5_0)+(n5_1))
denominatorx5 = sqrt(p.commonx5 * (1-p.commonx5) *
(1/(n5_0) + 1/(n5_1)))
z.prop.risx5 = numeratorx5
return(z.prop.risx5)}
zx5=z.propx5(x5_0,x5_1,n5_0,n5_1) #memanggil z score
p.value_x5=2*pnorm(-abs(zx5))

## test balance setelah strata untuk data continu __ X1
#strata 1
bx10_1= newdata$X1[newdata$strata == "1" &
newdata$X6 == "0"]
bx11_1= newdata$X1[newdata$strata == "1" &
newdata$X6 == "1"]
nx10_1=length(bx10_1)
nx11_1=length(bx11_1)
ttest <- function(m1,m2,s1,s2,n1,n2,m0=0,
equal.variance=FALSE)
{diff_mean <- (m1-m2-m0)
return(diff_mean)}
tx11=ttest( mean(bx10_1), mean(bx11_1), sd(bx10_1),
sd(bx11_1), nx10_1, nx11_1)
#strata 2
bx10_2= newdata$X1[newdata$strata == "2" &
newdata$X6 == "0"]
bx11_2= newdata$X1[newdata$strata == "2" &

```

```

newdata$X6 == "1"]
nx10_2=length(bx10_2)
nx11_2=length(bx11_2)
ttest <- function(m1,m2,s1,s2,n1,n2,m0=0,
equal.variance=FALSE)
{diff_mean <- (m1-m2-m0)
return(diff_mean)}
tx12=ttest( mean(bx10_2), mean(bx11_2), sd(bx10_2),
sd(bx11_2), nx10_2, nx11_2)
:
:
## test balance setelah strata untuk data kategori __ X5
#strata1
ax51_1=newdata$X5[(newdata$X5=="0" &
newdata$strata == "1" & newdata$X6 == "0")]
ax51_2=newdata$X5[(newdata$X5=="1" &
newdata$strata == "1" & newdata$X6 == "0")]
ax51_3=newdata$X5[(newdata$X5=="0" &
newdata$strata == "1" & newdata$X6 == "1")]
ax51_4=newdata$X5[(newdata$X5=="1" &
newdata$strata == "1" & newdata$X6 == "1")]
X5tes1 = as.table(rbind(c(length((ax51_1)),
length((ax51_2))),c(length((ax51_3)),
length((ax51_4))))
colnames(X5tes1) <- rownames(X5tes1) <- c("0", "1")
names(dimnames(X5tes1)) = c("X6", "X5")
x5_01=length((ax51_1))
n5_01=length((ax51_1))+length((ax51_2))
x5_11=length((ax51_3))
n5_11=length((ax51_3))+length((ax51_4))
z.propx51 = function(x5_01,x5_11,n5_01,n5_11){
Numeratorx51 = ((x5_01)/(n5_01)) - ((x5_11)/(n5_11))
p.commonx51 = ((x5_01)+(x5_11)) / ((n5_01)+(n5_11))
denominatorx51 = sqrt(p.commonx51 * (1-p.commonx51) *
(1/(n5_01) + 1/(n5_11)))
z.prop.risx51 = numeratorx51
return(z.prop.risx51)}
zx51=z.propx51(x5_01,x5_11,n5_01,n5_11) #call z score

#strata2
ax52_1=newdata$X5[(newdata$X5=="0" &

```



```

newdata$strata == "2" & newdata$X6 == "0") ]
ax52_2=newdata$X5[ (newdata$X5=="1" &
newdata$strata == "2" & newdata$X6 == "0") ]
ax52_3=newdata$X5[ (newdata$X5=="0" &
newdata$strata == "2" & newdata$X6 == "1") ]
ax52_4=newdata$X5[ (newdata$X5=="1" &
newdata$strata == "2" & newdata$X6 == "1") ]
X5tes2 = as.table(rbind(c(length((ax52_1)),
length((ax52_2))),c(length((ax52_3)),
length((ax52_4))))
colnames(X5tes2) <- rownames(X5tes2) <- c("0", "1")
names(dimnames(X5tes2)) = c("X6", "X5")
x5_02=length((ax52_1))
n5_02=length((ax52_1))+length((ax52_2))
x5_12=length((ax52_3))
n5_12=length((ax52_3))+length((ax52_4))
z.propx52 = function(x5_02,x5_12,n5_02,n5_12){
Numeratorx52 = ((x5_02)/(n5_02)) - ((x5_12)/(n5_12))
p.commonx52 = ((x5_02)+(x5_12)) / ((n5_02)+(n5_12))
denominatorx52 = sqrt(p.commonx52 * (1-p.commonx52) *
(1/(n5_02) + 1/(n5_12)))
z.prop.risx52 = numeratorx52
return(z.prop.risx52)}
zx52=z.propx52(x5_02,x5_12,n5_02,n5_12)#fungsi z score

#estimasi ATE
Y1.QT0 <- (newdata$Y[newdata$X8==0 &
newdata$strata==1])
Y1.QT1 <- (newdata$Y[newdata$X8==1 &
newdata$strata==1])
Y2.QT1 <- (newdata$Y[newdata$X8==1 &
newdata$strata==2])
Y2.QT0 <- (newdata$Y[newdata$X8==0 &
newdata$strata==2])
P11=mean(Y1.QT1)
P10=mean(Y1.QT0)
P21=mean(Y2.QT1)
P20=mean(Y2.QT0)

```

```

c(tx1,p.value_x2,p.value_x3,p.value_x4,tx5,p.value_x6,
p.value_x7,tx11,tx12,zx21,zx22,zx31,zx32,zx41,zx42,
tx51,tx52,zx61,zx62,zx61,zx72,P11,P10,P21,P20) }
> set.seed(2000)
> boot.bedamean2 <- boot(data_NDP, fn.boot.bedamean2,
R=200, stype="i")

```

Setelah bootstrap dilakukan, langkah selanjutnya adalah menghitung estimasi ATE dengan bootstrap pada masing-masing strata dan dilanjutkan dengan pengujian balance kovariat serta uji signifikansi parameter ATE yang sudah diestimasi. Berikut ini perintah yang diketikkan di R.

```

> Y11_bar=(boot.bedamean2$t[,16])
> Y10_bar=(boot.bedamean2$t[,17])
> Y21_bar=(boot.bedamean2$t[,18])
> Y20_bar=(boot.bedamean2$t[,19])
> mean_bootstrata1=Y11_bar-Y10_bar
> mean_bootstrata2=Y21_bar-Y20_bar
> theta1.boot=((nt1+nc1)*
  (mean_bootstrata1))/(nt+nc) #ate_S1
> theta2.boot=((nt2+nc2)*
  (mean_bootstrata2))/(nt+nc) # ate strata2
> theta_hat2.boot=(theta1.boot+theta2.boot) #ate boot
> theta1.bar.boot=mean(theta1.boot)
> theta2.bar.boot=mean(theta2.boot)
> theta_hat2.bar.boot=mean(theta_hat2.boot)
> num.se2=sum((theta_hat2.boot-theta_hat2.bar.boot)^2)
> denum.se2=boot.bedamean2$R - 1
> se2=num.se2/denum.se2 #estimasi SE bootstrap

## calculate p-value
> se_boot2=sqrt(diag(cov(boot.bedamean2$t))) # se
> z2_1.under.h0=(boot.bedamean2$t[,1]-
  boot.bedamean2$t0[1])/se_boot2[1]
> z2_2.under.h0=(boot.bedamean2$t[,2]-
  boot.bedamean2$t0[2])/se_boot2[2]
.
.
> z2_15.under.h0=(boot.bedamean2$t[,15]-

```

```

boot.bedamean2$t0[15])/se_boot2[15]
> z2_ate.under.h0=(theta_hat2 - theta_hat2.bar.boot)/se2

> p.value2_1 = 1/(1+boot.bedamean2$R) *
  (1+sum( ifelse(abs(z2_1.under.h0) >=
  abs(tx1$statistic),1,0)))
> p.value2_2 = 1/(1+boot.bedamean2$R) *
  (1+sum( ifelse(abs(z2_2.under.h0) >=
  abs(tx1$statistic),1,0)))
> p.value2_3 = 1/(1+boot.bedamean2$R) *
  (1+sum( ifelse(abs(z2_3.under.h0) >= abs(zx3),1,0)))
  .
  .
> p.value2_15 = 1/(1+boot.bedamean2$R) *
  (1+sum( ifelse(abs(z2_15.under.h0) >= abs(zx52),1,0)))
> p.value2_ate = 1/(1+boot.bedamean2$R) *
  (1+sum( ifelse(abs(z2_ate.under.h0) >= abs(t2),1,0)))

#summary balance bootstrap __ p value
> summary.bal.bootstrata2=cbind(c(p.value2_1,p.value2_6,
  p.value2_7),c(p.value2_2,p.value2_8,p.value2_9),
  c(p.value2_3,p.value2_10,p.value2_11),
  c(p.value2_4,p.value2_12,p.value2_13),
  c(p.value2_5,p.value2_14,p.value2_15))
> colnames(summary.bal.bootstrata2) <-
  c("X1", "X2","X3","X4","X5"

> rownames(summary.bal.bootstrata2) <-
  c("before strata","strata 1", "strata 2")

```

Diperoleh hasil pengujian balance dan estimasi ATE sebagai berikut

```

> summary.bal.bootstrata2

```

	X1	X2	X3	X4
before strata	0.5074627	0.00497512	0.0348259	0.5472637
strata 1	0.9054726	0.33333333	0.8557214	0.7462687
strata 2	0.4676617	0.46268657	0.2835821	0.8258706

```

                                X5
before strata 0.0049751
strata 1      0.1194030
strata 2      0.6218905

> theta1.bar.boot           # TE strata 1
[1] 0.2563184

> theta2.bar.boot           # TE strata 2
[1] 0.2047533

> theta_hat2.bar.boot       # ATE bootstrap
[1] 0.4610718

> se2                       # se bootstrap
[[1]] 0.01330783

> p.value2_ate              # p value ATE
[1] 0.004975124

```

Hasil pengujian balance dari output diatas menunjukkan bahwa sebelum dilakukan stratifikasi terdapat 3 kovariat yang *imbalance*, hal ini ditunjukkan oleh nilai *p-value* yang kurang dari $\alpha = 5\%$. Kovariat tersebut adalah lama pasien menderita DM (X2), jenis kelamin pasien (X3), dan status hipertensi (X5). Setelah dilakukan stratifikasi berdasarkan nilai propensity, diperoleh bahwa semua kovariat sudah balance pada semua strata ($p\text{-value} > \alpha = 5\%$). Hal ini menunjukkan bahwa metode propensity score stratification bootstrap telah mereduksi bias dari *confounding* dan tahapan bisa dilanjutkan ke estimasi ATE.

Berdasarkan output diatas juga diperoleh informasi bahwa efek perlakuan pada strata 1 sebesar 0,256 dan pada strata 2 sebesar 0,205, serta estimasi efek perlakuan (ATE) secara keseluruhan adalah sebesar 0,461 dengan standar error sebesar 0,013. Nilai $p\text{-value} < \alpha = 5\%$ menunjukkan bahwa kadar asam urat serum tinggi berpengaruh signifikan terhadap kejadian Neuropati Diabetik Perifer (NDP). Kesimpulan ini

sama dengan yang diperoleh ketika menggunakan PSS tanpa bootstrap.

References

- Akolo, I.R. *Propensity Score Stratification Bootstrap menggunakan Regresi Logistik Biner pada Kasus Diabetes Melitus (DM)*. Surabaya: Statistika ITS
- Austin, P.C. (2011). *A Tutorial and Case Study in Propensity Score Analysis: AN Application to Estimating the Effect of In-Hospital Smoking Cessation Counseling on Mortality*. *Multivariate Behavioral Research*, 46:119–151
- Dehejia, R.H., & Wahba, S. (1999). *Causal Effect in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs*. *Journal of the American Statistical Association*, Vol.94 No. 448
- 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
- Efron, B., & Tibshirani, Robert J. (1993). *An introduction to the Bootstrap*. New York: Chapman & Hall
- Guo, S. & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. California: Sage Publications
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics*, 86, pp. 4–29
- Islamiah, F. *Propensity Score menggunakan Regresi Logistik pada Kasus Data HIV/AIDS LSM Orbit Surabaya*. Surabaya: Statistika ITS
- 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), pp. 383-390.

- McCaffrey, D.F., Ridgeway, G., & Moral, A.R. (2004). *Propensity Score Estimation with Boosted Regression for Evaluating Causal Effect in Observational Studies*. *Psychological Method, 9(4)*, pp 403.
- McCandless, L.C., Gustafson, P., & Austin, P.C. (2009). *Bayesian propensity score analysis for observational data*. *Statistics in Medicine, 28*, pp 94-112.
- Ningsih, S. (2016).
- Pan, W., & Bai, H. (2015). *Propensity Score Analysis: Fundamental and Developments*. New York: Gulford Press
- Rosenbaum, P.R. (2010). *Design of Observational Studies*. New York: Springer
- Rosenbaum, P.R., & Rubin, D.B. (1983). *The Central Role of the Propensity Score in Observational Studies for Causal Effects*. *Journal Biometrika, vol.70, No.1*, pp 41-55.
- Tu, W., & Zhou, X.H. (2002). *A Bootstrap Confidence Interval Procedure for the Treatment Effect Using Propensity Score Subclassification*. *Health Services & Outcomes Research Methodology 3*: 135–147
- Yanovitzky, I., Zanutto, E., & Hornik, R. (2005). *Estimating Causal Effects of Public Health Education Campaigns using Propensity Score Methodology*. *Journal Elsevier Evaluation and Program Planning 28 (2005) pp. 209–220.*

Lampiran

1. Data DM untuk Propensity Score Stratification (PSS) dan PSS Bootstrap

No	Y	X1	X2	X3	X4	X5	X6
1	1	60	5.5	0	1	0	0
2	1	45	2.5	1	0	0	1
3	1	59	5	1	1	0	0
4	1	59	5	0	1	0	1
5	1	49	2	1	0	0	0
6	1	45	1.5	1	1	0	0
7	1	65	16	1	0	0	1
8	1	58	4	1	0	0	0
9	1	46	18	1	0	0	1
10	1	62	5	0	0	0	1
11	1	33	2.5	0	1	0	1
12	1	40	5	1	0	0	1
13	1	48	3	1	1	0	1
14	1	57	15	1	1	0	0
15	0	39	3	1	1	0	1
16	0	58	5	1	1	0	0
17	0	56	1.5	0	0	0	0
18	0	50	2	1	0	0	0
19	0	52	6	1	0	0	1
20	0	59	3	1	1	0	0
21	0	56	2	0	0	0	0
22	0	58	2.5	1	0	0	0
23	0	45	5	0	1	0	0
24	0	47	5	1	0	0	0
25	0	61	2	0	0	0	0
26	0	51	4	1	1	0	0
27	0	55	5	1	1	0	0
28	0	55	2.5	0	0	0	0

29	0	41	2.5	0	1	0	0
30	0	50	5	1	1	0	0
31	0	54	2	0	1	0	0
32	0	47	1.6	1	0	0	0
33	0	55	3	1	0	0	0
34	1	51	12	0	1	1	1
35	1	58	2.5	1	0	1	1
36	1	59	10	1	1	1	0
37	1	46	8	0	1	1	1
38	1	42	10	1	1	1	0
39	1	43	4	0	0	1	1
40	1	65	6	0	1	1	0
41	1	56	15	0	0	1	1
42	1	62	4	0	0	1	1
43	1	51	3	0	0	1	1
44	1	65	6	0	1	1	1
45	1	55	7	0	1	1	1
46	1	65	3.5	1	1	1	1
47	1	51	10	0	1	1	1
48	1	44	9	0	0	1	1
49	1	50	14	0	1	1	1
50	1	53	2	0	1	1	1
51	1	60	5	0	1	1	1
52	1	63	5	0	1	1	1
53	1	59	13	0	1	1	1
54	1	62	8	1	1	1	1
55	1	57	12	1	1	1	1
56	1	63	9	0	1	1	1
57	1	65	10	1	1	1	0
58	1	61	13	0	0	1	1
59	1	61	15	1	1	1	1
60	1	48	1.5	0	0	1	0
61	1	50	5	0	1	1	1
62	1	56	10	1	0	1	1
63	1	59	20	1	0	1	1

64	1	56	10	0	0	1	1
65	1	55	6	0	1	1	1
66	1	51	10	0	0	1	1
67	1	65	10	0	1	1	1
68	1	51	5	0	1	1	1
69	1	55	6	0	1	1	0
70	1	64	3	0	1	1	1
71	1	55	20	0	0	1	0
72	1	65	6	0	1	1	1
73	0	52	8	0	1	1	0
74	0	39	7	1	0	1	0
75	0	65	6	1	0	1	0
76	0	63	10	1	0	1	1
77	0	62	3.5	0	0	1	1
78	0	46	3	0	0	1	0
79	0	61	4	0	0	1	0
80	0	48	6	0	1	1	0
81	0	63	1.5	1	0	1	0
82	0	65	8	1	0	1	0

1. Data HIV LSM Surabaya untuk Propensity Score Weighting

No	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
1	0	1	37	2	2	3	1	4	0	2	1	1	1	2
2	0	1	31	2	2	3	1	2	0	1	2	1	2	2
3	0	1	30	2	2	2	1	2	0	4	1	1	2	2
4	0	1	43	2	2	3	3	4	0	1	1	2	2	2
5	1	1	31	2	1	4	3	4	0	4	3	1	2	2
6	0	1	37	2	1	4	3	4	0	1	1	2	2	2
7	0	1	33	2	1	4	1	2	1	1	1	1	2	2
8	1	2	28	1	2	2	1	1	1	4	1	2	2	2
9	1	2	28	1	2	2	1	1	0	4	1	2	2	2
10	0	2	23	1	1	4	1	4	0	4	3	2	2	2
11	0	1	32	0	1	4	3	4	0	1	1	2	1	2
12	0	1	29	0	1	4	1	2	0	4	3	2	2	2
13	1	1	34	1	2	4	3	4	0	3	1	2	2	2
14	0	1	42	0	1	4	1	2	1	1	2	2	2	1
15	0	1	32	2	1	4	1	2	0	4	3	2	2	2
16	0	1	34	2	1	4	3	4	0	1	1	1	2	2
17	0	1	46	2	1	4	3	4	0	1	1	2	2	2
18	0	1	35	2	1	4	3	4	0	4	2	1	1	2
19	0	1	35	2	1	4	3	2	0	4	1	2	1	2
20	1	1	38	2	1	4	3	4	0	4	1	1	1	2
21	0	1	31	2	1	4	1	4	0	4	2	1	1	2
22	0	1	32	2	1	4	3	4	0	4	2	2	2	1
23	0	1	32	0	1	4	3	4	0	1	1	1	1	2
24	0	1	35	2	2	2	3	4	0	4	1	1	1	2
25	1	1	35	2	1	3	3	4	0	1	3	2	2	2
26	0	1	37	1	1	4	3	2	0	4	1	1	1	2
27	0	1	39	2	2	3	1	4	0	1	2	1	1	2

28	0	1	30	2	2	4	1	2	0	4	1	2	2	2
29	1	1	24	0	1	4	3	4	0	4	3	1	2	2
30	0	1	35	0	1	4	3	4	0	3	1	2	2	2
31	0	1	35	2	2	3	1	1	0	4	1	1	1	2
32	0	1	31	0	1	4	3	4	0	4	2	1	1	2
33	0	1	32	1	2	2	3	4	0	4	1	2	2	2
34	0	1	34	2	1	4	3	4	0	4	1	1	1	2
35	0	1	30	2	1	4	3	4	0	1	2	1	1	2
36	0	1	30	2	1	4	3	4	0	4	1	1	1	2
37	0	1	31	0	1	4	3	4	0	4	1	1	2	2
38	0	1	32	0	1	4	3	4	0	1	1	1	2	2
39	0	1	31	2	2	1	1	3	0	4	1	2	2	1
40	0	1	29	2	1	4	3	4	0	4	1	1	1	2
41	0	1	33	2	1	4	3	2	0	1	1	1	1	2
42	0	1	35	2	2	4	3	2	0	4	2	2	2	1
43	1	1	35	3	1	3	3	4	0	1	1	2	2	1
44	0	1	37	2	2	3	3	4	0	4	1	2	1	2
45	1	1	22	0	1	4	3	2	0	4	3	1	2	2
46	0	1	39	2	1	4	1	2	0	4	1	2	2	2
47	0	1	32	2	2	3	3	4	0	1	1	1	2	2
48	0	1	35	2	1	4	3	4	0	1	1	2	2	2
49	0	1	32	1	2	4	3	4	0	4	3	2	1	1
50	0	1	37	2	1	4	3	4	0	4	3	2	2	2
51	0	1	37	1	2	4	3	4	0	4	2	2	2	2
52	0	1	41	2	1	4	3	4	1	1	1	1	1	1
53	0	1	36	2	2	2	3	2	0	4	1	2	2	2
54	0	1	34	0	1	4	1	4	0	4	3	2	2	2
55	0	1	34	2	1	4	1	2	0	4	1	2	2	2
56	0	1	33	0	1	4	3	4	0	4	3	2	2	2
57	0	1	36	1	1	4	1	4	0	1	1	1	1	2
58	0	1	36	1	1	4	1	4	0	4	1	1	1	2
59	0	1	27	2	1	4	3	4	0	4	1	1	2	2

60	1	1	29	3	2	3	3	2	1	4	1	2	2	2
61	1	1	35	2	2	2	3	4	0	4	2	2	2	1
62	0	1	33	2	2	3	1	2	0	1	1	1	1	2
63	0	1	25	1	2	2	1	2	0	4	1	1	2	2
64	0	1	48	2	2	1	3	4	0	1	1	2	1	2
65	0	1	31	2	2	2	3	4	0	1	1	1	1	1
66	0	1	28	2	2	2	3	2	0	4	1	1	2	2
67	0	1	33	2	1	4	2	4	0	1	2	1	2	2
68	0	1	38	2	1	4	1	4	0	4	2	2	2	2
69	1	1	32	2	1	4	3	4	1	4	2	1	1	2
70	0	1	35	2	1	4	1	2	0	4	2	2	2	1
71	0	1	31	2	1	4	3	4	0	1	1	1	1	2
72	0	1	28	2	2	2	3	2	0	1	3	2	2	2
73	0	1	30	2	1	4	1	2	0	4	1	1	1	1
74	1	2	41	0	1	4	3	4	0	1	3	2	2	2
75	0	1	38	2	2	2	1	2	0	4	3	2	1	2
76	1	1	33	2	1	4	3	2	1	4	1	2	2	1
77	1	1	33	0	1	4	3	4	0	4	2	2	2	2
78	1	1	30	2	1	4	3	4	0	4	2	1	1	2
79	1	1	30	2	1	4	3	2	0	4	3	2	2	1
80	0	1	26	2	1	4	3	4	0	1	1	1	1	2
81	0	1	34	0	1	4	3	4	0	1	2	1	1	2
82	0	1	33	2	1	4	2	4	1	1	2	1	1	1
83	0	1	23	2	2	2	1	2	0	4	3	2	2	2
84	0	1	25	2	1	4	1	4	0	1	2	1	1	2
85	0	1	33	0	1	4	3	4	0	1	1	2	2	2
86	0	1	35	2	2	4	3	4	0	4	3	2	2	2
87	0	1	31	2	2	3	3	4	0	1	1	1	1	1
88	0	1	24	2	1	4	1	2	0	1	2	2	1	2
89	0	2	24	2	2	3	2	4	0	2	1	1	2	2
90	0	1	32	2	2	2	1	2	0	4	1	2	2	2

91	0	1	28	2	2	2	1	2	0	4	1	1	2	2
92	1	1	26	2	2	4	3	2	0	1	2	2	2	2
93	1	1	28	2	2	4	3	4	0	4	1	1	2	2
94	0	1	36	0	1	4	3	4	0	4	2	1	1	2
95	0	1	33	0	1	4	3	4	0	4	1	1	1	2
96	1	1	28	2	1	4	3	4	0	4	1	2	2	2
97	0	1	37	0	1	4	3	4	0	4	2	2	1	1
98	0	1	49	2	2	3	3	4	0	4	1	2	2	2
99	0	1	28	3	1	4	1	4	0	4	2	2	2	2
100	1	1	23	1	2	4	1	2	0	2	2	2	2	2
101	0	1	36	3	2	3	1	4	0	3	2	2	2	1
102	0	1	35	2	1	2	3	4	1	4	3	1	1	1
103	0	1	33	2	2	2	3	4	0	5	3	2	2	1
104	1	1	31	2	1	4	3	1	0	4	2	2	2	1
105	1	1	29	2	2	2	1	2	0	4	1	2	2	1
106	0	1	30	2	1	4	3	4	0	4	1	1	2	2
107	0	1	35	2	2	4	1	2	0	1	1	2	2	1
108	0	1	33	2	2	2	1	4	0	4	2	1	1	1
109	0	1	33	2	1	4	1	2	1	1	1	1	1	2
110	0	1	33	2	1	4	1	2	0	1	1	1	1	2
111	0	1	31	0	1	4	1	2	0	3	1	2	1	2
112	1	1	28	2	1	4	1	2	0	4	3	1	1	2
113	0	1	33	2	1	4	3	4	0	4	1	1	2	2
114	1	1	26	0	1	4	1	4	0	4	1	2	2	2
115	0	1	37	1	1	4	3	2	0	1	1	1	1	2
116	0	1	37	2	1	4	3	4	0	1	1	2	2	2
117	0	1	31	2	1	4	3	4	0	4	1	1	2	2
118	1	1	33	2	1	4	3	2	0	5	1	2	2	2
119	1	1	32	2	1	4	3	4	0	1	1	1	1	1
120	0	1	34	0	1	4	3	4	0	1	1	1	1	2
121	0	1	28	3	2	3	3	4	1	1	1	1	2	1
122	0	1	31	2	1	4	3	4	0	4	1	1	1	2

123	1	1	38	1	2	4	3	4	0	4	1	2	1	2
124	0	1	36	2	1	4	1	2	0	4	2	1	1	2
125	0	1	34	2	1	4	3	4	0	4	2	1	1	1
126	1	1	25	1	1	4	1	4	0	4	1	1	1	2
127	0	1	35	0	1	4	3	4	0	4	2	2	2	2
128	0	1	30	3	2	3	3	3	0	1	3	1	1	1
129	0	1	33	1	1	4	3	2	0	4	1	2	2	2
130	0	1	32	2	1	4	3	4	0	1	1	1	1	2
131	0	1	46	2	2	3	1	4	0	2	2	1	1	2
132	0	1	31	2	1	4	3	2	0	4	1	2	2	2
133	1	1	35	1	2	2	3	4	0	4	3	2	2	1
134	0	1	37	0	1	4	3	4	0	1	3	2	2	2
135	0	1	37	1	2	4	1	2	0	4	3	2	2	1
136	0	1	28	0	1	4	1	2	0	3	1	1	2	2
137	0	1	22	1	1	4	3	4	0	4	1	1	2	2
138	1	2	29	3	2	3	1	3	0	1	1	1	1	2
139	0	1	29	2	2	4	1	2	0	4	1	1	1	2
140	1	1	28	2	1	4	1	2	0	4	1	2	2	2
141	0	1	40	2	2	2	1	2	0	4	1	1	1	2
142	0	1	42	2	2	3	3	4	0	1	1	2	1	2
143	0	1	27	0	1	4	2	4	0	4	1	1	1	2
144	0	1	32	2	2	2	3	4	0	4	1	2	2	2
145	0	1	31	2	2	3	3	4	0	1	1	1	2	2
146	0	1	29	2	2	3	1	2	0	1	2	1	1	2
147	0	1	33	2	2	2	3	4	0	4	1	1	2	2
148	1	1	25	2	1	4	1	4	1	1	2	1	1	1
149	0	1	28	2	1	4	1	2	0	4	1	2	2	2
150	1	1	58	0	1	4	3	4	0	1	1	1	1	1
151	1	1	33	0	1	4	3	4	0	4	1	1	1	2
152	0	1	46	2	1	4	1	2	0	4	1	2	2	2
153	1	1	32	2	2	3	3	4	0	4	1	2	2	2

154	0	1	35	2	2	3	3	4	0	1	1	1	1	2
155	0	1	18	0	1	4	1	2	0	2	1	1	2	2
156	0	1	36	0	1	4	3	4	0	1	1	1	1	2
157	1	1	33	2	2	3	3	4	0	1	2	2	2	2
158	0	1	28	2	2	2	3	4	1	4	2	1	1	1
159	0	1	29	2	2	2	3	4	0	4	1	2	2	2
160	0	1	34	2	1	4	1	4	0	4	2	2	2	2
161	0	1	28	2	2	3	3	4	0	4	1	2	1	2
162	1	1	31	2	1	4	3	4	0	4	1	1	1	2
163	1	1	22	0	1	4	1	2	0	1	2	1	1	2
164	0	1	36	0	1	4	1	2	0	4	2	2	2	2
165	0	1	28	2	1	4	3	4	0	4	1	2	2	2
166	1	1	35	2	1	4	3	4	0	4	1	1	2	2
167	0	1	34	1	1	4	3	4	0	1	1	1	2	2
168	0	1	31	2	2	3	3	4	0	4	1	1	1	2
169	0	1	31	2	2	3	3	4	0	4	1	1	1	2
170	0	1	32	2	2	2	1	4	0	4	3	2	2	2
171	0	1	43	3	1	4	3	4	0	1	1	1	1	2
172	0	1	26	2	1	4	3	2	0	4	2	2	2	2
173	0	1	33	2	2	3	3	4	0	4	1	1	2	2
174	1	1	29	0	1	4	1	2	0	1	2	1	1	2
175	0	1	24	0	1	4	1	2	0	4	3	2	2	2
176	1	1	33	2	1	4	1	2	0	4	2	2	2	1
177	0	1	33	2	1	4	1	2	0	4	1	1	2	2
178	0	1	34	2	1	4	1	2	0	3	1	1	2	2
179	1	1	35	2	2	2	1	2	0	4	1	2	2	2
180	0	1	37	2	1	4	3	4	0	4	1	2	2	2
181	0	1	38	2	1	4	1	4	0	4	2	1	1	2
182	0	1	34	2	1	4	1	2	0	4	1	2	2	2
183	0	1	29	1	1	4	1	3	0	1	3	2	2	2
184	0	1	31	2	2	2	3	4	1	4	3	1	1	2
185	0	1	36	2	2	2	3	4	1	5	1	1	1	1

186	0	1	34	2	1	4	1	2	0	4	1	2	2	2
187	0	1	29	2	2	3	1	4	0	1	2	1	1	2
188	0	1	31	0	1	4	3	4	0	1	1	1	1	2
189	0	1	34	2	1	4	1	2	0	4	3	1	1	2
190	0	1	33	0	1	4	3	2	0	4	1	2	2	2
191	0	2	38	2	1	4	2	4	0	1	1	2	2	2
192	0	1	34	0	1	4	3	4	0	1	1	2	2	2
193	0	1	34	0	1	4	3	4	1	1	1	1	1	2
194	0	1	30	1	1	4	3	4	0	4	3	2	2	2
195	0	1	28	0	1	4	3	4	0	2	1	1	1	2
196	0	1	47	2	1	2	3	4	0	4	2	1	1	2
197	0	1	38	2	2	3	3	4	0	1	1	1	2	2
198	0	1	38	2	2	3	3	4	0	1	1	1	1	1
199	0	1	25	2	1	4	3	4	0	4	3	1	2	2
200	0	1	29	0	1	4	3	4	0	4	1	1	1	2
201	1	1	35	2	2	3	3	4	0	4	3	2	1	2
202	0	1	34	2	1	4	1	4	0	4	1	1	2	2
203	1	1	27	2	1	4	1	2	0	1	1	2	2	2
204	0	1	32	2	1	4	1	4	0	1	1	1	1	2
205	0	1	33	0	1	4	3	4	0	4	3	2	2	2
206	0	1	33	0	1	4	3	4	0	4	2	2	2	2
207	0	1	29	2	1	4	1	2	0	4	1	2	2	2
208	0	1	34	2	1	4	3	4	0	4	2	1	2	2
209	1	1	35	0	1	4	3	4	1	1	1	1	1	1
210	0	1	35	0	1	4	3	4	0	4	1	1	1	2
211	0	1	33	2	1	4	1	2	0	4	2	1	1	2
212	0	1	34	2	2	4	1	4	0	1	1	2	2	2
213	0	1	39	2	1	4	3	4	0	4	1	2	2	2
214	0	1	26	2	1	4	3	4	0	4	1	2	2	2
215	0	1	39	2	1	4	3	2	0	4	3	2	2	2
216	0	1	23	0	1	4	1	3	0	3	1	1	2	2

217	1	1	33	1	1	4	3	4	0	4	1	1	2	2
218	1	1	31	1	1	4	1	4	0	4	1	1	1	1

2. Data HIV RS X untuk Propensity Score with Marginal Mean Weighting through Stratification

RES	IO (Y)	X1.1	x1.2	X1.3	X1.4	X1.5	X2.1	X2.2	KPH©
1	2	1	1	1	2	2	3	2	1
2	2	4	1	3	2	2	2	3	2
3	1	1	2	1	1	1	2	1	1
4	2	3	1	2	2	2	3	2	2
5	2	1	1	2	1	1	3	3	2
6	2	1	1	2	1	1	3	3	1
7	1	1	1	1	2	2	1	1	1
8	1	1	1	1	1	1	2	2	1
9	2	1	1	3	2	1	3	3	1
10	2	2	1	2	2	2	3	2	2
11	2	1	1	2	1	1	4	3	2
12	2	2	1	3	2	2	2	3	2
13	1	4	1	3	2	1	1	1	2
14	1	4	2	3	2	1	2	2	2
15	1	2	1	2	2	2	3	2	1
16	1	1	1	1	1	1	2	2	1
17	1	1	2	1	1	1	1	1	1
18	1	1	2	2	1	1	2	2	1
19	2	2	1	2	2	2	4	3	2
20	2	2	1	2	2	2	3	2	1
21	1	2	1	2	1	1	1	1	1
22	2	1	1	2	2	2	3	2	1
23	1	1	2	1	1	2	1	1	1
24	2	2	1	2	2	2	4	3	2
25	2	4	1	3	2	2	2	3	1
26	1	1	1	2	1	1	2	2	1

58	2	1	2	2	1	1	3	2	2
59	1	4	2	3	2	2	2	2	1
60	2	1	1	2	1	1	4	3	2
61	2	1	2	2	2	1	3	3	2
62	1	1	1	1	1	2	2	2	1
63	2	1	2	2	1	2	3	3	2
64	1	1	2	1	2	1	2	2	2
65	1	1	1	1	1	1	2	2	1
66	1	1	1	1	1	1	3	3	1
67	2	1	2	2	1	2	4	3	2
68	1	2	1	1	2	1	2	2	1
69	1	2	2	1	2	2	2	2	2
70	1	1	1	1	2	2	1	1	2
71	1	2	1	1	2	1	2	2	1
72	2	1	2	2	1	1	3	2	2
73	1	1	2	1	1	2	3	2	2
74	1	1	2	1	1	1	2	2	1
75	2	1	1	1	1	1	4	3	2
76	1	1	1	2	2	2	3	2	2
77	2	4	2	3	2	2	4	3	2
78	1	1	2	1	1	1	2	2	1
79	1	1	1	2	1	1	2	2	2
80	1	1	2	1	1	1	3	2	1
81	2	2	2	1	2	2	3	3	1
82	2	1	2	2	2	1	2	2	2
83	1	1	2	1	1	1	2	2	2
84	2	2	2	1	2	2	3	2	2
85	1	1	2	2	1	1	1	1	1
86	2	2	1	1	2	2	3	3	2
87	1	1	2	1	1	1	2	2	1
88	1	1	2	1	1	1	1	1	1
89	2	1	2	1	1	1	4	3	2

RIWAYAT PENULIS



Dr. Bambang Widjanarko Otok, M.Si lahir di Surabaya pada tahun 1968. Lulusan S1 Program studi Statistika, Institut Teknologi Sepuluh November (ITS) Surabaya pada tahun 1994, lulus pendidikan Magister Statistika di Institut Pertanian Bogor pada tahun 2000 serta program doktoral bidang Statistika Universitas Gajah Mada pada tahun 2008. Saat ini beliau memiliki jabatan Lektor Kepala. Beliau bekerja sebagai dosen mulai tahun 1994. Sampai sekarang beliau berprofesi sebagai dosen dan Kepala Laboratorium Statistika Bisnis dan Sosial di Institut Teknologi Sepuluh November Surabaya.

Beliau aktif melakukan riset dan pengajaran pada beberapa bidang statistika seperti Teori Statistika, Komputasi Statistika, Teknik Survei dan *Sampling*, Model Linier, Analisis Data, Analisis Multivariate dan lain-lain. Beliau juga aktif sebagai narasumber pelatihan statistik seperti Pemodelan statistik untuk industry, Analisis Regresi Linier, SEM, AMOSS, Pemodelan Time Series, dan lain-lain. Sejumlah karya tulis dipublikasikan dalam jurnal internasional bereputasi seperti *European Journal of Scientific Research*, *Statistic in Science*, dan nasional terakreditasi sebagai bentuk kepedulian beliau dalam penyebaran ilmu pengetahuan.



Dr. Puhadi, M.Sc lahir di Sidoarjo, pada tanggal 04 Februari 1962. Lulus dari pendidikan S1 Matematika Institute Teknologi Bandung pada tahun 1986 beliau mengabdikan sebagai dosen di ITS Surabaya. Lulus S2 Matematika Statistika dari Universitas Waseda Tokyo pada tahun 1992 sedangkan program doctoral di Universitas Gajah Mada Yogyakarta bidang keilmuan Matematika Statistika pada tahun 2005 dengan judul disertasi Rancangan Percobaan Optimal Dan Kondisi Optimum pada Model Permukaan Multirespon.

Beliau mengampu mata kuliah Kalkulus, Probabilitas, Matematik Statistik, Analisis Data Kualitatif, Model Linier dan Statistik Inferensi. Dalam dua tahun terakhir beliau memenangkan program Hibah Penelitian Dasar Unggulan PT Kemenristekdikti dengan judul Penentuan Standar Pengendalian Pelayanan Minimal Prevalensi Ispa dan Diare Suatu Wilayah dengan Pendekatan Spasial Mars dan Generalized Bivariate Poisson Regression, serta pada tahun 2018 mampu kembali menerima program Hibah Kompetensi Kemenristekdikti Penaksiran Parameter Dan Pengujian Hipotesis Pada Geographically Weighted Trivariate Generalized Poisson Regression (Studi Kasus : Jumlah Kematian Bayi, Balita dan Kematian Ibu di Provinsi Jawa Timur Tahun 2017).

Karya tulis beliau banyak dipublikasikan dalam jurnal internasional bereputasi diantaranya Applied Mathematical Sciences dengan judul Parameter Estimation of Geographically Weigthed Multivariate Poisson Regression (2015), Parameter Estimation Of Geographically Weighted Multivariate T Regression Model dalam Journal of Theoretical and Applied Information Technology (2016), dan lain-lain.



Dr. Mahdalena, S.Pd., MKes merupakan dosen Poltekkes Kementerian Kesehatan Banjarmasin yang bertugas sebagai Kepala Pusat Penelitian dan Pengabdian Masyarakat Poltekkes Kemenkes Banjarmasin. Beliau tinggal di Kompleks Banua Permai No.344 RT. 39 Rw. 07 Kelurahan Sungai Besar Banjarbaru – 70714. Lulus pendidikan Diploma 3 Keperawatan di Akper Depkes

Banjarmasin pada tahun 1992, Program Bidan B di Akper Pajajaran pada tahun 1994, Program Sarjana Pendidikan di Universitas Islam Kalimantan Muhammad Arsyad Al Banjary pada tahun 1999, Magister Ilmu Kesehatan Masyarakat Universitas Airlangga Surabaya pada tahun 2002 serta lulus dalam program doktoral Ilmu Kesehatan Masyarakat di Universitas Airlangga pada tahun 2017. Sebagai dosen dan ketua Lemlit yang menghasilkan banyak karya penelitian dan pengabdian masyarakat maka telah banyak artikel dalam jurnal terindeks scopus yang telah beliau hasilkan sebagai bentuk publikasi seluruh kegiatan akademik yang dilaksanakan.

PROPENSITY SCORE Untuk Data Kesehatan

Metode *Propensity score* didefinisikan sebagai probabilitas bersyarat untuk menerima suatu *treatment* berdasarkan pada karakteristik sebelum *treatment*. Buku ini membahas tentang konsep dasar dan metode *propensity score*, penggunaan *propensity score* dalam *stratification*, *weighting* serta dalam *bootstrap*.

Buku ini disusun dalam 6 bab yang diterjemahkan dari beberapa referensi dan memuat panduan praktis yang bisa diaplikasikan untuk data kesehatan. Buku ini dilengkapi lampiran-lampiran data yang bisa dipraktekkan pembaca.

Penerbit:
STIKes Majapahit Mojokerto
Jalan Raya Jabon KM 02 Mojoanyar
Mojokerto
Telp. 0321 329915
Fax. 0321 329915
Email: mojokertostikesmajapahit@gmail.com

