Forum Statistika dan Komputasi, Oktoberl 2005, p: 17 – 24 Vol. 10 No. 2
ISSN : 0853-8115

KLASIFIKASI SKOR PROPENSITAS DALAM PENDUGAAN SELANG KEPERCAYAAN BOOTSTRAP UNTUK PERBEDAAN NILAI TENGAH DUA POPULASI

Marzuki

Jurusan Matematika FMIPA UNSYIAH

Asep Saefuddin dan Anang Kurnia

Departemen Statistika FMIPA IPB

Abstract

The comparison of mean of two populations assumes that there is no other variables influence (covariate) except the difference of the observed variable. In real data, this condition is often unfulfilled. Propensity score classification (PSC) is a method to overcome the case. In this research, we do simulation data to evaluate the method, and as the illustration we do the real data of first semester NMR of IPB postgraduate students in Statistics major. The simulation data is generated by covariates either with the same means or different ones to both groups, each with different parameter (δ_0) 0,00, 0,25 and 0,50. The bootstrap confidence interval included a distribution which is built from propensity score estimations without the variance estimation.

The result shows that 95% bootstrap confidence interval with PSC method includes the parameter for the same and different covariate distribution respectively as 0,95 and 0,87. This method is suitable only when the sample sizes are larger. The illustration uses real data with covariate age, marital status, graduate (S-1) NMR and occupation as a lecturer or not, the result estimation of 95% bootstrap interval confidence to differentiate NMR of postgraduate Statistics students in IPB between those who came from the universities in Java and outside Java is between -0,13 and 0,72.

Keywords : propensity, bootstrap, covariates

PENDAHULUAN

Dalam suatu penelitian, untuk mengetahui perbedaan nilai tengah suatu peubah acak dari dua populasi dibandingkan antarcontoh dari masing-masing kelompok atau populasi tersebut. Apabila populasi melibatkan kovariat maka hal itu akan sulit dilakukan karena sebaran kovariat berbeda-beda. Misalkan akan menduga perbedaan Nilai Mutu Rataan (NMR) antara dua populasi mahasiswa pascasarjana, maka pendugaan itu akan sulit karena setiap individu mahasiswa mempunyai latar belakang yang berbeda seperti umur, status perkawinan, dan status pekerjaan yang diduga mempengaruhi parameter.

Dalam analisis regresi, untuk memperhitungkan perbedaan nilai tengah dua kelompok dapat ditempuh dengan memasukkan satu peubah bebas yaitu peubah indikator atau peubah *dummy* Z ke dalam model, sedangkan latar belakang mahasiswa seperti kasus di atas dapat dimasukkan ke dalam model sebagai peubah bebas lainnya. Jika Z=1 untuk kelompok yang satu dan Z=0 untuk kelompok yang lain maka harapan perbedaan nilai tengah antara dua kelompok tersebut direpresentasikan oleh koefisien dari Z, γ. Suatu hipotesis nol bahwa tidak ada perbedaan nilai tengah antara dua populasi dapat dengan mudah diperlihatkan melalui regresi dan uji hipotesis terhadap γ. Adapun kesimpulan yang diambil berdasarkan asumsi semua peubah lain konstan.

Salah satu metode yang digunakan untuk mengendalikan perbedaan pengaruh kovariat adalah dengan cara melibatkan pengelompokan unit perlakuan ke dalam kelas-kelas berdasarkan karakteristik yang diamati. Rosenbaum dan Rubin (1984) menyatakan bahwa salah satu karakteristik melalui klasifikasi skor propensitas (KSP) telah terbukti menjadi salah satu cara yang efektif

ISSN : 0853-8115

untuk mengurangi bias dalam menduga perbedaan nilai tengah dua populasi atau pengaruh perlakuan. Namun demikian respon di dalam tiap-tiap kelas dan di antara kelas tidak saling bebas karena klasifikasinya didasarkan pada urutan penduga skor propensitas yang umum diduga melalui model logistik. Sementara itu penduga dari skor propensitas yang tidak diketahui juga menambah sumber keragaman yang dapat mempengaruhi penduga ragam yang digunakan dalam mengambil kesimpulan (Tu *et al.* 1999, diacu dalam Tu & Zhou 2003) sehingga dalam pendugaan perbedaan nilai tengah dua populasi yang melibatkan kovariat dengan menggunakan KSP itu validitas prosedur tersebut di atas agak diragukan (Du 1998, diacu dalam Tu & Zhou 2003). Oleh karena itu, Tu dan Zhou (2003) memperkenalkan suatu metode KSP dalam pendugaan selang kepercayaan perbedaan nilai tengah dua populasi dengan bootstrap nonparametrik.

Dalam paper ini akan diurai suatu prosedur pendugaan perbedaan nilai tengah dua populasi dengan menggunakan KSP dan pendugaan selang kepercayaan bootstrapnya serta mengevaluasi metode KSP menggunakan data simulasi dengan menghitung peluang cakupan parameter dalam selang kepercayaan. Sebagai ilustrasi akan dilakukan pendugaan perbedaan NMR mahasiswa S-2 Program Studi Statistika Sekolah Pascasarjana Institut Pertanian Bogor (IPB) antara yang berasal dari lulusan S-1 di Jawa dan di luar Jawa.

LANDASAN TEORI

Klasifikasi Skor Propensitas

Metode skor propensitas diberlakukan untuk dua kelompok secara serentak. Setiap subjek individu yang menunjukkan indikator kelompok diberi notasi Z dengan Z=1 untuk kelompok ke-1 dan Z=0 untuk kelompok ke-2, sedangkan vektor kovariat yang terdapat dalam kelompok diberi notasi **X**. Peubah respon untuk setiap subjek dinotasikan oleh Y(Z), di mana $Y(Z=1)$ menunjukkan peubah respon dari subjek dalam kelompok ke-1, dan $Y(Z=0)$ menunjukkan peubah respon dari subjek dalam kelompok ke-2.

Gagasan awal dari metode ini adalah menggantikan koleksi dari kovariat-kovariat yang membaur dalam studi penelitian dengan satu fungsi dari kovariat-kovariat ini yang disebut skor propensitas (Rubin 1997). Nilai skor propensitas didefinisikan oleh Tu dan Zhou (2003) sebagai

$$e(\mathbf{X}) = P(Z=1 \mid \mathbf{X})$$

dan Sianesi (2001) menulis $0< e(\mathbf{X}) = P(Z=1 \mid \mathbf{X}=\mathbf{x}) <1$ untuk setiap $\mathbf{x} \in \mathbf{X}$ yang tidak lain adalah nilai peluang bersyarat suatu kelompok berdasarkan kovariat-kovariat yang diamati. Nilai ini dapat digunakan untuk mengendalikan bias karena ketidakseimbangan kovariat yang diamati. Parsons (2001) menyatakan KSP merupakan salah satu cara untuk mengurangi bias dari suatu kelompok yang melibatkan kovariat.

Harapan dari perbedaan nilai tengah dua kelompok dinyatakan oleh

$$\delta = E[Y(1) - Y(0)]$$

Hanya ada satu dari dua hasil ($Y(1)$ dan $Y(0)$) yang teramati dalam menduga δ sehingga tidak bisa langsung menduga δ berdasarkan data hasil pengamatan (Rosenbaum & Rubin 1984). Salah satu cara mengatasi masalah tersebut adalah penggunaan skor propensitas $e(\mathbf{X})$.

Rosenbaum dan Rubin (1983), diacu dalam Tu dan Zhou (2003) menunjukkan bahwa

$$\delta = E_{e(\mathbf{X})}\ [E\{Y(1) \mid e(\mathbf{X}),Z=1\} - E\{Y(0) \mid e(\mathbf{X}),Z=0\}],$$

sedangkan $E_{e(\mathbf{X})}$ menunjukkan harapan dengan mengacu pada sebaran dari $e(\mathbf{X})$ dalam populasi subjek. Jika subjek bisa dibagi menjadi K kelas I_1, …, I_K yang mempunyai lebar kelas yang sama berdasarkan skor propensitas dan misalkan bahwa paling sedikit ada satu subjek dalam setiap kelas yang berasal dari setiap populasi, maka akan didapat

$$\delta = \sum_{k=1}^{K} P(e(\mathbf{X}) \in I_k)[E\{Y(1) \mid e(\mathbf{X}) \in I_k, Z=1\} - E\{Y(0) \mid e(\mathbf{X}) \in I_k, Z=0\}]$$

Pendugaan skor propensitas $e(\mathbf{X})$ dilakukan melalui model logistik

$$\ln\left(\frac{P(Z=1 \mid \mathbf{X}=\mathbf{x})}{1-P(Z=1 \mid \mathbf{X}=\mathbf{x})}\right) = \mathbf{x}^t\beta$$

sehingga penduga skor propensitas $e(\mathbf{X})$ adalah

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

Semua subjek distratifikasi menjadi K kelas dengan menggunakan penduga skor propensitas sehingga setiap subjek dalam setiap kelas mempunyai nilai penduga skor propensitas yang hampir sama. Misalkan n_{1k} dan Y_{1k1}, …, $Y_{1kn_{1k}}$ adalah jumlah subjek dan respon dalam kelompok ke-1 dalam kelas ke-k. n_{0k} dan Y_{0k1}, …, $Y_{0kn_{ok}}$ adalah jumlah subjek dan respon dalam kelompok ke-2 dalam kelas ke-k; dengan k = 1, …, K. Nilai tengah respon untuk kelompok ke-1 dan kelompok ke-2 dalam kelas ke-k dinyatakan oleh

$$\overline{Y}_{1k} = \sum_{i=1}^{n_{1k}} Y_{1ki} / n_{1k} \text{ dan } \overline{Y}_{0k} = \sum_{i=1}^{n_{0k}} Y_{0ki} / n_{0k}$$

Jumlah total subjek dalam kelompok ke-1 dan kelompok ke-2 adalah

$$n_1 = \sum_{k=1}^{K} n_{1k} \text{ dan } n_0 = \sum_{k=1}^{K} n_{0k}$$

Persamaan (2) memperlihatkan bahwa untuk menduga δ perlu menduga proporsi skor

ISSN : 0853-8115

propensitas dalam setiap kelas, $P(e(\mathbf{X})\in I_k)$ dan nilai harapan dari peubah respon kedua kelompok dalam setiap kelas, $E(Y(1)\,|\,e(\mathbf{X})\in I_k, Z=1)$ serta $E(Y(0)\,|\,e(\mathbf{X})\in I_k, Z=0)$. Ketiga parameter tersebut diduga oleh masing-masing statistik berikut

$$P(e(\mathbf{X})\in I_k) \,\hat{=}\, (n_{1k}+n_{0k})/(n_1+n_0)$$

$$E(Y(1)\,|\,e(\mathbf{X})\in I_k, Z=1) \,\hat{=}\, \overline{Y}_{1k}$$

$$E(Y(0)\,|\,e(\mathbf{X})\in I_k, Z=0) \,\hat{=}\, \overline{Y}_{0k}$$

yang kemudian digunakan untuk menduga perbedaan nilai tengah dua populasi

$$\hat{\delta} = \sum_{k=1}^{K}\left(\frac{(n_{1k}+n_{0k})}{n_1+n_0}(\overline{Y}_{1k}-\overline{Y}_{0k})\right) \quad (5)$$

Statistik $\hat{\delta}$ yang diberikan dalam persamaan (5) menggunakan titik-titik potong yang ditentukan oleh kuantil dari penduga skor propensitas kedua kelompok.

Pendugaan Selang Kepercayaan Bootstrap

Bootstrap merupakan suatu metode simulasi berdasarkan pada data untuk menarik kesimpulan secara statistika. Simulasi bootstrap di antaranya berguna untuk menduga ukuran akurasi suatu statistik dari suatu contoh tunggal, menduga bias suatu statistik dari suatu contoh dan membentuk dugaan selang kepercayaan (Efron & Tibshirani 1993). Salah satu prosedur penduga selang kepercayaan dengan metode bootstrap nonparametrik berdasarkan bootstrap *Bias-Corrected accelerated* (*BCa*). Prosedur tersebut diawali dengan mencocokkan model propensitas yang ada pada persamaan (3) menggunakan data contoh ($\mathbf{X}_i$, Z_i) untuk $i = 1,2,\ldots,n$. Setelah pencocokan model, penduga skor propensitas ($\hat{e}_i$) dihitung untuk setiap subjek menggunakan persamaan (4). Berdasarkan pada skor propensitas ini, subjek dipartisi menjadi K kelas yang mempunyai lebar kelas yang sama. Penduga perbedaan nilai tengah dua populasi $\hat{\delta}$ dihitung dengan menggunakan persamaan (5).

Setiap iterasi bootstrap b, $b = 1, 2, \ldots, B$, dilakukan penarikan contoh dengan pengembalian terhadap semua subjek dalam kelompok ke-1 (n_1) dan semua subjek yang ada dalam kelompok ke-2 (n_0) secara terpisah dari data contoh asli (bukan data contoh bootstrap). Misalkan $(Y_{i'}^{(b)}, \mathbf{X}_{i'}^{(b)}, Z_{i'}^{(b)})$ adalah data contoh bootstrap ke-b, $i' = 1, 2, \ldots, n$. Data hasil penarikan contoh pengembalian $(\mathbf{X}_{i'}^{(b)}, Z_{i'}^{(b)})$, untuk $i = 1, 2, \ldots, n$, dicocokkan kembali menggunakan model logistik dan penduga skor propensitas dihitung untuk setiap subjek dalam contoh bootstrap, $\hat{e}_{i'}^{(b)}$, kemudian peubah respon bootstrap $Y_{i'}^{(b)}$ distratifikasi dan selanjutnya dihitung perbedaan nilai tengah dua populasi $\hat{\delta}^{(b)}$.

Setelah melakukan proses di atas sebanyak B kali, perhitungan penduga selang kepercayaan bootstrap BCa 100(1-α)% untuk perbedaan nilai tengah dua populasi δ dapat dilakukan dengan langkah-langkah sebagai berikut. Efron dan Tibshirani (1993) menyatakan konstanta koreksi bias dihitung dengan menggunakan persamaan

$$\hat{d} = \Phi^{-1}\left(\frac{\#[\hat{\delta}^{(b)} < \hat{\delta}]}{B}\right)$$

sedangkan $\Phi(\cdot)$ adalah fungsi sebaran kumulatif dari sebaran normal standar. Selanjutnya dalam Efron dan Tibshirani (1993) parameter akselerasi untuk dua contoh dihitung dengan menggunakan persamaan

$$\hat{a} = \frac{1}{6}\hat{\sigma}^{-3/2}(n_1^{-2}\hat{\gamma}_1 - n_0^{-2}\hat{\gamma}_0)$$

dengan

$$\hat{\sigma}^2 = \hat{\sigma}^2_{1jack}n_1^{-1} + \hat{\sigma}^2_{0jack}n_0^{-1}$$

$$\hat{\sigma}^2_{1jack} = \left[\frac{n_1-1}{n_1}\sum_{i=1}^{n_1}(s^2_{1(i)} - \bar{s}^2_{1(.)})^2\right]$$

$$\hat{\sigma}^2_{0jack} = \left[\frac{n_0-1}{n_0}\sum_{i=1}^{n_0}(s^2_{0(i)} - \bar{s}^2_{0(\cdot)})^2\right]$$

$$\bar{s}^2_{1(\cdot)} = \sum_{i=1}^{n_1}s^2_{1(i)}/n_1$$

$$\bar{s}^2_{0(\cdot)} = \sum_{i=1}^{n_0}s^2_{0(i)}/n_0$$

di mana $s^2_{1(i)}$ dan $s^2_{0(i)}$ masing-masing adalah ragam data kelompok ke-1 dan kelompok ke-2 untuk data yang pengamatan ke-i dihilangkan. Persamaan (9) dan (10) adalah penduga ragam *jackknife* dari contoh asli untuk menduga ragam yang tidak diketahui dari kedua kelompok.

Selanjutnya dihitung kemiringan contoh dari masing-masing kelompok

$$\hat{\gamma}_1 = \frac{\sum_{i=1}^{n_1}(Y_{1i}-\overline{Y}_1)^3/n_1}{\left[\sum_{i=1}^{n_1}(Y_{1i}-\overline{Y}_1)^2/n_1\right]^{3/2}}$$

dan

$$\hat{\gamma}_0 = \frac{\sum_{i=1}^{n_0}(Y_{0i}-\overline{Y}_0)^3/n_0}{\left[\sum_{i=1}^{n_0}(Y_{0i}-\overline{Y}_0)^2/n_0\right]^{3/2}}$$

dengan

$\overline{Y}_1 = \sum_{i=1}^{n_1}Y_{1i}/n_1$ dan $\overline{Y}_0 = \sum_{i=1}^{n_0}Y_{0i}/n_0$

Penduga perbedaan nilai tengah dua populasi bootstrap diurutkan dari nilai yang terkecil ke nilai yang terbesar,

$\hat{\delta}^{(1)} \le \hat{\delta}^{(2)} \le \ldots \le \hat{\delta}^{(B)}$. Penduga selang kepercayaan bootstrap *BCa* 100(1-α)% untuk perbedaan nilai tengah dua populasi didefinisikan oleh Tu dan Zhou (2003) sebagai

$$\left(\hat{\delta}^{[B(\hat{\beta}_{a(\alpha/2)})]}, \hat{\delta}^{[B(\hat{\beta}_{a(1-\alpha/2)})]}\right) \quad (13)$$

dengan

$$\hat{\beta}_{a(\alpha/2)} = \Phi\left\{\hat{d} + (\hat{d} + z_{\alpha/2})\{1 - \hat{a}(\hat{d} + z_{\alpha/2})\}^{-1}\right\}$$

$$\hat{\beta}_{a(1-\alpha/2)} = \Phi\left\{\hat{d} + (\hat{d} + z_{1-\alpha/2})\{1 - \hat{a}(\hat{d} + z_{1-\alpha/2})\}^{-1}\right\}$$

dan [*x*] adalah nilai integer terkecil yang lebih besar atau sama dengan *x*.

KAJIAN SIMULASI

Simulasi dilakukan dengan terlebih dahulu membangkitkan kovariat-kovariat X. Pada simulasi ini digunakan empat kovariat. Pembangkitan dilakukan dengan dua cara yaitu kovariat yang mempunyai sebaran sama dan kovariat yang mempunyai sebaran berbeda antardua kelompok.

Pembangkitan untuk kovariat yang mempunyai sebaran sama diasumsikan bahwa X_1~N(μ_1, 1), X_2~N(μ_2, 1), X_3~N(μ_3, 1) dan X_4~N(μ_4, 1) dengan jumlah data bangkitan 1000 masing-masing 500 untuk kelompok ke-1 dan 500 untuk kelompok ke-2. Adapun untuk cara pembangkitan kedua, agar sebaran kovariat berbeda secara sistematik dalam dua kelompok maka dibangkitkan kovariat-kovariat untuk dua kelompok secara terpisah, yaitu untuk kelompok ke-1 (Z=1) diasumsikan bahwa X_1~N(μ_{11}, 1,5), X_2~N(μ_{11}, 1,5), X_3~N(μ_{21}, 1,5) dan X_4~N(μ_{21}, 1,5), sedangkan untuk kelompok ke-2 (Z=0) diasumsikan bahwa X_1~N(μ_{10}, 1,5), X_2~N(μ_{10}, 1,5), X_3~N(μ_{20}, 1,5) dan X_4~N(μ_{20}, 1,5). Masing-masing kelompok dibangkitkan data sebanyak 500 dengan *setting* parameter seperti disajikan pada Tabel 1. Kajian simulasi dilakukan sebanyak 100 kali untuk setiap *setting*.

Tabel 1. *Setting* parameter untuk simulasi

Setting	δ_0	μ_1	μ_2	μ_3	μ_4
1	0,00	0	0,5	1	1,5
2	0,25	0	0,5	1	1,5
3	0,50	0	0,5	1	1,5
	δ_0	μ_{10}	μ_{20}	μ_{11}	μ_{21}
4	0,00	0	1	1	1,5
5	0,25	0	1	1	1,5
6	0,50	0	1	1	1,5

Berdasarkan kedua cara bangkitan itu kemudian dibangkitkan respon Y dari suatu hubungan linier Y=Zδ_0+$\mathbf{X}^t\boldsymbol{\beta}$+ε berdasarkan data kovariat-kovariat dan indikator kelompok di atas dengan menggunakan nilai $\boldsymbol{\beta}$=(β_1 β_2 β_3 β_4)′=(0,5 0,4 0,5 0,4)′ dan membangkitkan secara bebas galat normal ε~N(0, 1) untuk masing-masing δ_0=0,00, 0,25, 0,50 sedangkan δ_0 menunjukkan perbedaan nilai tengah dua populasi yang ditetapkan untuk membangkitkan data respon Y. Pembangkitan data ini dilakukuan dengan perangkat lunak MATLAB 6.5.

Semua kovariat dalam simulasi ini dibangkitkan dari sebaran normal untuk memudahkan dalam mengevaluasi metode yang digunakan. Sebenarnya kovariat X tidak ditentukan sebarannya karena pendugaan skor propensitas yang berdasarkan vektor kovariat tersebut dilakukan melalui model logistik yang mengasumsikan bahwa peubah bebas (dalam hal ini adalah kovariat X) ditetapkan (*fixed*).

- **Metode t-student**

Data yang dibangkitkan dengan cara pertama (kovariat bersebaran sama) memperlihatkan bahwa dengan menggunakan metode ini maka penduga selang kepercayaan 95% perbedaan nilai tengah dari kedua kelompok untuk δ_0=0,00 dan δ_0=0,25 berpeluang mencakup parameter masing-masing sebesar 0,95. Penduga selang kepercayaan 95% dari perbedaan nilai tengah dari kedua kelompok untuk δ_0=0,50 berpeluang mencakup parameter sebesar 0,94. Berdasarkan data yang dibangkitkan dengan cara kedua (kovariat bersebaran berbeda) diperoleh bahwa penduga selang kepercayaan 95% dari perbedaan nilai tengah untuk δ_0=0,00 δ_0=0,25 dan δ_0=0,50 mempunyai peluang sebesar 0,00.

- **Metode model regresi**

Pendugaan selang kepercayaan 95% perbedaan nilai tengah kedua kelompok dengan menggunakan metode model regresi diperoleh bahwa untuk data bangkitan cara pertama, selang mencakup parameter untuk δ_0=0,00, δ_0=0,25 dan δ_0=0,50 mempunyai peluang masing-masing 0,99, 0,96, dan 0,92. Data bangkitan cara kedua, selang mencakup parameter untuk δ_0=0,00, δ_0=0,25 dan δ_0=0,50 mempunyai peluang masing-masing 0,94, 0,94, dan 0,93.

- **Metode klasifikasi skor propensitas**

Selang kepercayaan untuk metode KSP diduga dengan menggunakan selang kepercayaan

ISSN : 0853-8115

bootstrap. Nilai penduga selang kepercayaan bootstrap untuk perbedaan nilai tengah dua kelompok yang diperoleh memperlihatkan bahwa untuk data bangkitan cara pertama, selang mencakup parameter untuk δ_0=0,00, δ_0=0,25 dan δ_0=0,50 mempunyai peluang masing-masing sebesar 0,96, 0,97 dan 0,93. Adapun untuk data bangkitan cara kedua, selang mencakup parameter untuk δ_0=0,00, δ_0=0,25 dan δ_0=0,50 mempunyai peluang masing-masing sebesar 0,89, 0,83, dan 0,90.

Nilai tengah perbedaan $\bar{\delta}$ dan peluang cakupan paramater perbedaan dua populasi yang ditetapkan yaitu δ_0 dalam selang kepercayaan yang diduga untuk ke tiga jenis metode dan ke enam *setting* disajikan pada Tabel 2. Adapun P menyatakan peluang cakupan paramater perbedaan nilai tengah dua populasi yang ditetapkan.

Tabel 2. Nilai tengah perbedaan dan peluang cakupan parameter

Setting	t-*student*		Regresi		KSP	
	$\bar{\delta}$	P	$\bar{\delta}$	P	$\bar{\delta}$	P
1	-0,003	0,95	0,005	0,99	0,001	0,96
2	0,241	0,95	0,244	0,96	0,240	0,97
3	0,475	0,94	0,496	0,92	0,490	0,93
4	1,376	0,00	0,095	0,94	-0,021	0,89
5	1,628	0,00	0,229	0,94	0,342	0,83
6	1,8541	0,00	0,465	0,93	0,575	0,90

CONTOH PENERAPAN

Data riil yang digunakan untuk ilustrasi berasal dari data mahasiswa S-2 Program Studi Statistika IPB tahun 1999 sampai dengan 2004 dengan jumlah pengamatan 101 mahasiswa. Data tersebut menggambarkan mahasiswa S-2 Statistika IPB dengan karakteristik yang diamati terdiri dari kovariat-kovariat X_1, X_2, X_3 dan X_4 berturut-turut sebagai usia (tahun), status (menikah atau tidak), NMR S-1 dan pekerjaan (dosen atau tidak). Peubah respon (Y) yang menjadi perhatian adalah NMR semester ke-1, sedangkan yang menjadi indikator kelompok (Z) adalah lulusan S-1, dengan Z=1 adalah lulusan S-1 di Pulau Jawa dan Z=0 adalah lulusan S-1 di luar Pulau Jawa.

Jumlah pengamatan yang berstatus menikah pada saat semester ke-1 sebanyak 50,50% dan yang bekerja sebagai dosen sebanyak 69,31%, sedangkan rataan usianya adalah 29,95 tahun dan rataan NMR S-1 adalah 3,0477. Adapun rataan NMR semester ke-1 sebagai peubah respon adalah 3,2431 dengan simpangan baku 0,4607. Ringkasan statistik dari data mahasiswa tersebut disajikan dalam Tabel 3 dan Tabel 4.

Penduga perbedaan nilai tengah dari NMR semester ke-1 antara kelompok Jawa dan luar Jawa adalah sebesar 0,121 dengan selang kepercayaan 95% berdasarkan t-*student* adalah (-0,081; 0,323). Artinya tidak ada perbedaan yang nyata antara kedua kelompok tersebut.

Evaluasi menggunakan model regresi menghasilkan dugaan model sebagai berikut

$E[Y]$ = 2,05 - 0,00929 X1 + 0,167 X2 + 0,299 X3 + 0,391 X4 + 0,280 Z

dengan koefisien determinasi R^2 = 27,6%. Jadi nilai harapan dari NMR semester ke-1 untuk mahasiswa lulusan S-1 di Pulau Jawa dan di luar Pulau Jawa masing-masing diduga oleh

$E[Y \mid X_1, X_2, X_3, X_4, Z=1]$ = 2,05 - 0,00929 X1 + 0,167 X2 + 0,299 X3 + 0,391 X4 + 0,280

$E[Y \mid X_1, X_2, X_3, X_4, Z=0]$ = 2,05 - 0,00929 X1 + 0,167 X2 + 0,299 X3 + 0,391 X4

sehingga

$E[Y \mid X_1, X_2, X_3, X_4, Z=1] - E[Y \mid X_1, X_2, X_3, X_4, Z=0]$ = 0,280 yang secara statistik berbeda nyata dengan

Tabel 3. Deskripsi data riil dari peubah respon dan kovariat kontinu

	Y (NMR sem ke-1)	X_1 (usia)	X_3 (NMR S-1)
Nilai tengah	3,2431	29,950	3,0477
Simpangan baku	0,4607	6,540	0,4607
Median	3,2500	28,000	3,0100
Minimum	2,0000	22,000	2,2900
Maksimum	4,0000	54,000	3,9200

Tabel 4. Deskripsi data riil dari kelompok dan kovariat biner

		Jumlah mahasiswa
Z (lulusan)	Z=1 (Jawa)	75
	Z=0 (luar Jawa)	26
X_2 (status)	X_2 =1 (menikah)	51
	X_2 =0 (belum menikah)	50
X_4 (pekerjaan)	X_4 =1 (dosen)	70
	X_4 =0 (bukan dosen)	31

nilai p = 0,005. Artinya perbedaan nilai harapan dari NMR semester ke-1 antara kelompok Jawa dan luar Jawa adalah sebesar 0,280 dengan asumsi keempat kovariatnya konstan.

Evaluasi dengan KSP dilakukan dengan menduga nilai skor propensitas sehingga diperoleh wilayah sebesar 0,5955. Tabel 5 diperoleh dengan membagi subjek pengamatan menjadi 3 kelas. Pembagian subjek pengamatan ke dalam 3 kelas dilakukan agar terpenuhi asumsi bahwa minimal ada satu subjek dalam setiap kelas yang berasal dari setiap populasi.

Berdasarkan Tabel 5 maka penduga perbedaan NMR semester ke-1 antara mahasiswa S-2 Program Studi Statistika Sekolah Pascasarjana IPB yang berasal dari lulusan S-1 di Jawa dan di luar Jawa, dengan menggunakan persamaan (9) adalah 0,376087.

Tabel 5. Nilai-nilai hasil partisi data riil

Kelas (k)	Batas Kelas	n_{1k}	n_{0k}	$\bar{Y}_{1k}$	$\bar{Y}_{0k}$
1	0,376 ≤ I_1 ≤ 0,574	9	6	3.456	3.333
2	0,574 < I_2 ≤ 0,773	24	18	3.429	3.157
3	0,773 < I_3 ≤ 0,971	42	2	3.147	2.585
Jumlah		75	26		

Tabel 6. Statistik penduga perbedaan contoh bootstrap data riil

Nilai tengah	0,34003
Simpangan baku	0,21458
Median	0,34473
Minimum	-0,44291
Maksimum	0,85107

Nilai penduga selang kepercayaan bootstrap untuk perbedaan nilai tengah dua populasi dihitung sehingga diperoleh statistik penduga perbedaan contoh bootstrap serta nilai penduga ragam dan kemiringan dua populasi yang disajikan pada Tabel 6 dan Tabel 7.

Banyaknya penduga perbedaan nilai tengah untuk contoh bootstrap yang lebih kecil dari penduga perbedaan nilai tengah untuk contoh asli adalah 558 sehingga dengan mengunakan persamaan (10) diperoleh konstanta koreksi bias $\hat{d}$ = 0,145900. Nilai penduga ragam dan kemiringan untuk data riil pada Tabel 7 memberikan nilai parameter akselerasi $\hat{a}$ = -0,125008. Berdasarkan konstanta koreksi bias dan nilai parameter akselerasi di atas maka $\hat{\beta}_{a(\alpha/2)}$ = 0,0138966 dan $\hat{\beta}_{a(1-\alpha/2)}$ = 0,965078 sehingga urutan batas bawah selang adalah 14 dan urutan batas atas selang adalah 966. Jadi penduga selang kepercayaan bootstrap perbedaan nilai tengah dua populasi adalah (-0,133191; 0,717376).

Ringkasan penduga perbedaan nilai tengah dua populasi dan selang kepercayaannya untuk ketiga jenis metode disajikan pada Tabel 8.

DISKUSI

Metode KSP yang diterapkan dalam pendugaan selang kepercayaan bootstrap untuk perbedaan nilai tengah dua populasi yang dikemukakan oleh Tu dan Zhou (2003) tidak memerlukan penduga ragam dan cocok diterapkan dalam penelitian dengan ukuran contoh yang besar. Metode ini tidak berdasarkan pada asumsi sebaran apapun yang membatasi kovariat-kovariat yang terlibat. Selang kepercayaan bootstrap yang dihasilkan meliputi sebaran yang dibangun dari penduga skor propensitas dan mengakomodasikan ketergantungan di antara respon di dalam dan antarkelas sekaligus dalam kaitannya dengan pembentukan struktur yang diawali dengan pengkelasan. Sehingga prosedur ini dapat diterapkan relatif mudah bila melibatkan banyak kovariat.

Hasil simulasi memberikan gambaran apabila suatu respon melibatkan kovariat yang mempunyai sebaran yang berbeda untuk dua

Tabel 7. Nilai penduga ragam dan kemiringan dua populasi data riil

$\hat{\sigma}^2_{1jack}$	$\hat{\sigma}^2_{0jack}$	$\hat{\sigma}^2$	$\hat{\gamma}_1$	$\hat{\gamma}_0$
0,00103769	0,00181361	0,000083590	-0,351231	0,401043

Tabel 8. Ringkasan penduga perbedaan dan selang kepercayaan data riil

t-student		Model regresi		Skor propensitas	
$\hat{\delta}$	SK	$\hat{\delta}$	SK	$\hat{\delta}$	SK
0,121	(-0,081; 0,323)	0,280	(0,086; 0,474)	0,376	(-0,133; 0,717)

ISSN : 0853-8115

kelompok maka metode t-*student* tidak baik digunakan dalam menduga perbedaan nilai tengahnya karena hasilnya akan berbias. Hal ini ditunjukkan dengan tidak satupun selang kepercayaan yang mencakup parameter untuk ketiga *setting* 4, 5 dan 6.

Selang kepercayaan 95% dengan metode model regresi mencakup parameter untuk semua *setting* berpeluang di atas 92%. Hal ini menunjukkan bahwa jika respon bebar-benar merupakan fungsi dari kovariat maka model regresi baik digunakan. Akan tetapi masalahnya respon belum tentu dapat difungsikan dengan baik oleh kovariat. Untuk mengukur kebaikan model bisa digunakan R^2 atau R^2(adj) dengan batas tertentu.

Selang kepercayaan 95% dengan metode KSP untuk kovariat yang sama mencakup parameter δ_0 berpeluang di atas 93%. Hal ini menunjukkan bahwa metode KSP juga baik digunakan untuk menduga perbedaan nilai tengah dua populasi yang mempunyai sebaran kovariat sama. Adapun untuk sebaran kovariat yang berbeda, selang kepercayaan 95% bootstrap mencakup parameter berpeluang 89%, 83% dan 90% untuk masing-masing *setting* simulasi 4, 5 dan 6. Dengan demikian apabila suatu respon melibatkan kovariat, walaupun memiliki sebaran yang berbeda, maka metode KSP dengan selang kepercayaan bootstrap cocok digunakan untuk menduga selang kepercayaan perbedaan nilai tengah dua populasi. Namun demikian, masih perlu kajian sehubungan dengan menurunnya tingkat ketelitian selang kepercayaan yang dibentuk terutama pada saat kovariat berbeda.

Metode KSP untuk data riil memberikan penduga selang kepercayaan bootstrap untuk perbedaan nilai tengah NMR semester ke-1 antara mahasiswa S-2 lulusan S-1 dari Jawa dan luar Jawa adalah antara -0,133 dan 0,717. Selang kepercayaan ini mengandung nilai nol yang menunjukkan tidak ada perbedaan yang nyata antara dua populasi tersebut. Hal ini mungkin disebabkan pengelompokan perguruan tinggi di Jawa dan luar Jawa akan menghasilkan pengelompokan yang lebih beragam di dalam kelompok dibandingkan antar kelompok. Masing-masing kelompok (Jawa dan luar Jawa) terdapat perguruan tinggi yang berstatus negeri dan yang berstatus swasta dengan perkembangan dan kualitas yang berbeda-beda.

PENUTUP

KSP mampu menduga selang kepercayaan bagi perbedaan nilai tengah dua populasi dengan baik untuk kondisi kovariat yang berbeda maupun tidak berbeda. Pemilihan taraf nyata α=0,05 memberikan peluang cakupan parameter sebesar 83% hingga 90% untuk kondisi kovariat yang berbeda. Ini sedikit menurun jika dibandingkan dengan kondisi kovariat sama yang memberikan peluang cakupan parameter sebesar 93% hingga 97%.

Kajian terhadap data riil, selang kepercayaan bootstrap untuk perbedaan nilai tengah NMR semester ke-1 mahasiswa S-2 Program Studi Statistika Sekolah Pascasarjana IPB tahun 1999 sampai dengan 2004 adalah (-0,133; 0,717) yang berarti perbedaan tersebut sesungguhnya tidak nyata karena selang mengandung nilai nol.

Sehubungan dengan hal tersebut di atas perlu diperhatikan hal-hal sebagai berikut:

1. Kajian dapat dilakukan untuk mengevaluasi dan memperbaiki penurunan taraf nyata α sehingga memberikan peluang cakupan parameter yang lebih baik terutama untuk kondisi sebaran kovariat yang berbeda antar dua kelompok.
2. Pengelompokan perguruan tinggi Jawa dan luar Jawa yang selama ini sering dilakukan, agaknya kurang tepat karena keragaman di dalam kelompok masih relatif besar dibandingkan antarkelompok. Pengelompokan harusnya juga mempertimbangkan antara perguruan tinggi negeri dan swasta bahkan perguruan tinggi yang berstatus BHMN dan non-BHMN agar keragaman di dalam kelompok dapat diperkecil sehingga akan diperoleh hasil-hasil kajian yang lebih baik.

DAFTAR PUSTAKA

Efron B, Tibshirani RJ. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York

Parsons LS. 2001. Reducing Bias in a Propensity Score Matched-Pair Sample Using Greedy Matching Techniques. *Ovation Research Group* paper 214-26

Rosenbaum PR, Rubin DB. 1983. Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70:41-55

Rosenbaum PR, Rubin DB. 1984. Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association* 79:318-328

Rubin DB. 1997. Estimation from Nonrandomized Treatment Comparisons Using Subclassification on Propensity Scores. *Nonrandomized Comparative Clinical Studies* pp. 757-763

Tu W, Zhou XH. 2003. A Bootstrap Confidence Interval Procedure for the Treatment Effect Using Propensity Score Subclassification. *UW Biostatistics Working Paper Series* paper 200.

⊗