

PEMILIHAN UJI NONPARAMETRIK TERBAIK UNTUK DUA SAMPEL BEBAS MELALUI METODE SIMULASI

Sugiyanto¹, Etik Zukhronah²

^{1,2}Jurusan Matematika FMIPA-UNS,

e-mail : ¹Sugiy50@yahoo.co.id, ²etikzukhronah@yahoo.co.id

Abstrak

Telah tersedia beberapa uji nonparametrik, seperti uji Cramer-von Mises, uji Kolmogorov-Smirnos, uji Wilcoxon, dan uji peringkat baru yang dapat digunakan untuk menentukan apakah dua sampel tersebut berasal dari populasi yang sama atau tidak. Secara analitik keempat uji tersebut tidak dapat dibandingkan, namun demikian dalam makalah ini secara simulasi dapat dibandingkan dalam rangka untuk mencari uji terbaik diantara uji-uji tersebut.

Berdasarkan hasil simulasi menunjukkan bahwa uji peringkat baru mempunyai kepekaan tertinggi didalam menolak H_0 bila ternyata H_0 salah. Ini berarti uji peringkat baru merupakan uji terbaik di antara uji-uji tersebut.

Kata kunci : dua sampel bebas, uji Wilcoxon, uji Kolmogorov-Smirnov, uji Cramer von Mises, dan uji peringkat baru, uji terbaik.

PENDAHULUAN

Untuk mengetahui apakah dua sampel bebas berasal dari populasi yang sama, telah tersedia empat uji nonparametrik, yaitu uji Cramer von mises (1931), uji Kolmogorov-Smirnov (1933), uji Wilcoxon (1945), dan uji peringkat baru (1998).

Apabila diambil hipotesis null H_0 : menyatakan dua sampel berasal dari populasi yang sama, sedangkan hipotesis alternatif H_a : menyatakan dua sampel bukan berasal dari populasi yang sama, maka keempat uji tersebut dapat dipakai. Secara analitik keempat uji tersebut tidak dapat dibandingkan, namun demikian melalui metode simulasi akan dibandingkan guna memilih uji yang paling peka menolak H_0 ketika H_0 salah.

Perbedaan menolak H_0 apabila H_0 salah dari Uji Cramer-von Mises, Uji Kolmogorov-Smirnov, dan Uji Wilcoxon tersebut kecil (Conover, 1990). Pada tahun 1998, Baumgartner dkk. memperkenalkan Uji Peringkat Baru untuk dua sampel bebas. Uji Peringkat Baru ini merupakan pengembangan dari Uji Cramer von-Mises. Sugiyanto dan Rini. I. (2004) melalui simulasi berhasil menunjukkan adanya indikasi (gambaran beberapa kasus) bahwa Uji Peringkat Baru lebih peka menolak H_0 apabila H_0 salah (lebih kuasa) dibandingkan Uji Cramer-von Mises. Sugiyanto dan Ngapiningsih (2004) melalui simulasi berhasil menunjukkan adanya indikasi bahwa Uji Peringkat Baru lebih peka menolak H_0 apabila H_0 salah dibandingkan uji Kolmogorov-Smirnov. Sugiyanto dan Wulan. S (2005) melalui simulasi berhasil menunjukkan adanya indikasi bahwa Uji Peringkat Baru lebih peka menolak H_0 apabila H_0 salah dibandingkan Uji Wilcoxon. Sugiyanto dan Hasih P (2006) melalui simulasi berhasil menunjukkan adanya indikasi bahwa Uji Peringkat Baru lebih peka menolak H_0 apabila H_0 salah dibandingkan Uji Cramer-von Mises, Uji Kolmogorov-Smirnov dan Uji Wilcoxon. Sugiyanto dan Isnandar S (2007) berhasil menunjukkan bahwa uji t (uji parametrik) lebih peka menolak H_0 apabila H_0 salah dibandingkan Uji Peringkat Baru (uji nonparametrik). Menurut

Shabri dan Jemain (2008), belum ada kesepakatan ahli statistik untuk menentukan statistik uji yang paling kuat diantara uji Cramer von Mises, uji Kolmogorov-Smirnov, dan uji Wilcoxon. Sedangkan Ross dan Adam (2012) melalui simulasi Monte Carlo, menunjukkan uji Cramer von Mises lebih kuat dibandingkan uji Kolmogorov-Smirnov. Didalam penelitian ini akan dilakukan perbandingan dari empat uji tersebut guna memilih uji terbaik (mempunyai kepekaan tertinggi untuk menolak H_0 apabila H_0 salah).

PEMBAHASAN

Uji Cramer von Mises termasuk dalam kelompok statistik uji yang menggunakan fungsi distribusi empiris. Statistik uji Cramer von Mises dibentuk oleh kuadrat selisih distribusi masing-masing sampel $[F_m(x) - F_n(x)]^2$ yang diboboti oleh $(m+n)/mn$, dengan m menyatakan banyak sampel I, dan n menyatakan banyak sampel II. Uji Kolmogorov-Smirnov menggunakan jarak terjauh antara dua sampel $D_{m,n} = \max_x |F_m(x) - F_n(x)|$ sebagai statistik ujinya. Uji Wilcoxon menggunakan peringkat dari dua sampel yaitu $W_X = \sum_{i=1}^N iS_i$ sebagai statistik ujinya. Sedangkan uji peringkat baru menggunakan statistik uji $B = \frac{1}{2}(B_x + B_y)$ dengan

$$B_x = \frac{1}{n} \sum_{i=1}^n \frac{1}{i \left(1 - \frac{1}{n+1}\right)} \frac{\left(G_i - \frac{m+n}{n} \cdot i\right)^2}{\frac{(m+n)m}{n}},$$

dan

$$B_y = \frac{1}{m} \sum_{j=1}^m \frac{1}{j \left(1 - \frac{1}{m+1}\right)} \frac{\left(H_j - \frac{m+n}{m} \cdot j\right)^2}{\frac{(m+n)n}{m}}.$$

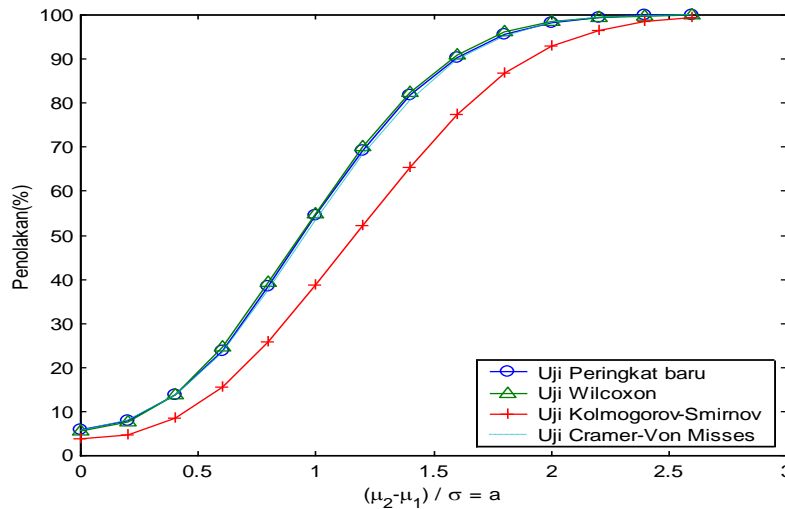
Metode simulasi merupakan salah satu cara yang sekarang ini sedang berkembang. Kemampuan menjangkau hal-hal yang lebih luas akan menjadi lebih jelas karena hanya membutuhkan asumsi lebih sedikit sehingga dapat digunakan untuk masalah yang lebih kompleks dibandingkan metode analitik. Menurut Efron (1990), simulasi adalah teknik berdasarkan angka untuk menuntuk suatu percobaandalam komputer digital dengan melibatkan matematika dan model logika yang menggambarkan kelakuan manusia, porses atau sistem.

Suatu pendekatan akan mendekati nilai sebenarnya jika pengulangan dalam simulasi cukup besar (Zhang, 2001). Oleh karena itu akan dilakukan pengulangan sebanyak 10000 kali. Penolakan H_0 bila H_0 salah dari setiap kali uji statistik diestimasi dari 10000 sampel random yang disimulasikan (Steele dan chaseling, 2006).

Perhatikan gambar 1, yang menggambarkan hasil simulasi dari dua populasi masing-masing berdistribusi normal dengan variansi sama tetapi mean membesar, yang diulang sebanyak 10.000 kali.

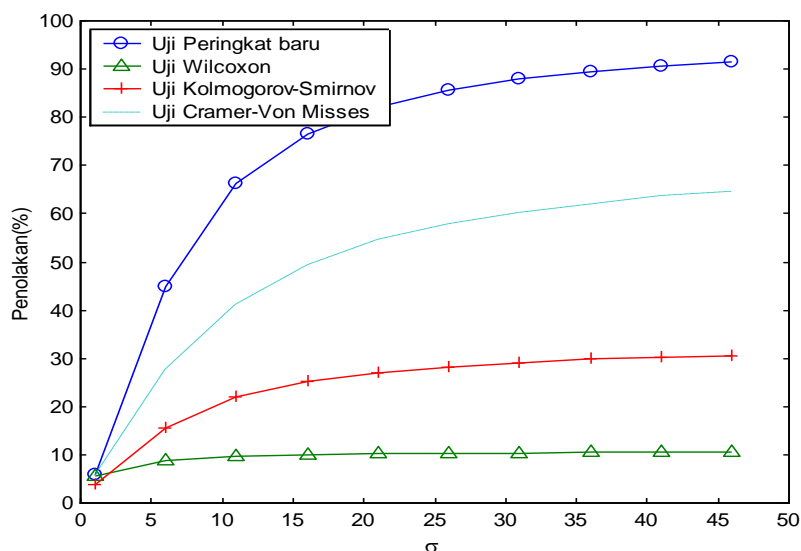
Ketika dua sampel yang dibangkitkan dari populasi yang sama dengan mean dan variansi sama pula, tampak keempat uji (uji Cramer von Mises, uji Kolmogorov-Smirnov, uji Wilcoxon, dan uji peringkat baru) tersebut mempunyai kepekaan yang sama untuk menolak H_0 ketika H_0 salah. Hasil ini sesuai dengan desain teori bahwa dua sampel yang berasal dari populasi yang sama (mempunyai karakteristik sama) akan menghasilkan penolakan yang yang amat kecil sesuai dengan tingkat

signifikansinya. Namun demikian dengan semakin besar perbedaan mean tampak uji Kolmogorov-Smirnov mempunyai kepekaan terendah untuk menolak H_0 ketika H_0 salah, akan tetapi ketika perbedaan mean mendekati 2,5 atau lebih hasil dari empat uji memberikan kesimpulan yang sama.



Gambar 1

Berdasarkan gambar tersebut dapat ditarik kesimpulan bahwa keempat uji mempunyai kuasa uji yang sama bila 2 sampel yang dibangkitkan berasal dari populasi yang sama dan mempunyai variansi yang sama pula, namun hanya berbeda meannya. Dalam hal ini bisa dikatakan bahwa keempat uji memberikan keputusan yang sama walau nilai mean berbeda.



Gambar 2

Hasil simulasi yang kedua, dapat dilihat pada gambar 2, dimana dua sampel yang dibangkitkan berasal dari distribusi normal dengan mean yang sama tetapi variansi berbeda-beda.

Sebagaimana yang telah diduga sebelumnya untuk mean dan variansi yang sama tampak bahwa empat uji menghasilkan keputusan yang sama. Sedangkan untuk perbedaan variansi, keputusan dari empat uji sangat berbeda. Hasil simulasi kedua ini memberikan gambaran yang sangat berbeda dengan hasil simulasi pertama. Tampak perbedaan keputusan diantara empat uji yang bersangkutan didalam penolakan H_0 ketika H_0 salah. Semakin besar perbedaan variansi semakin besar pula penolakan H_0 ketika H_0 salah.

Pada uji Kolmogorov-Smirnov dan uji Wilcoxon sangat rendah sekali penolakan terhadap H_0 ketika H_0 salah, bahkan uji Wilcoxon tidak mampu lagi untuk membedakan perbedaan dari dua sampel yang berbeda karakteristiknya utamanya yang mempunyai variansi berbeda. Tidaklah cukup untuk mengatakan bahwa dua sampel yang dibandingkan berasal dari populasi yang sama.

Apabila ditemukan suatu sampel yang mirip dengan sampel yang dibandingkan pada simulasi yang kedua maka pemakaian uji Cramer von Mises, uji Kolmogorov-Smirnov, dan uji Wilcoxon bisa menyesatkan karena berdasarkan hasil simulasi tampak ketiga uji tersebut sangat rendah sekali untuk menolak H_0 ketika H_0 salah.

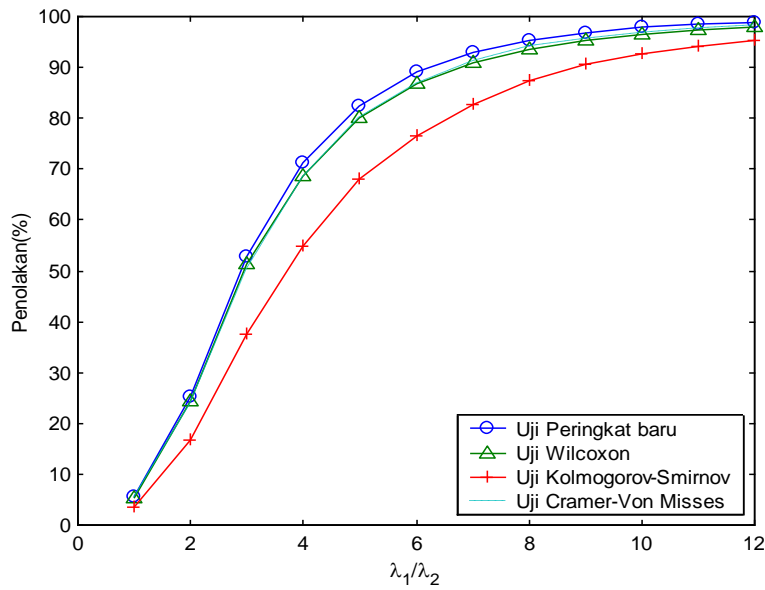
Hasil simulasi ketiga tampak pada gambar 3, yang mana dua sampel yang dibandingkan berasal dari populasi yang mempunyai distribusi eksponensial dengan perbandingan nilai parameter yang berbeda.

Ketika diambil perbandingan nilai parameter sama, ini berarti kedua sampel berasal dari populasi yang sama, maka hasil uji dari empat uji menghasilkan keputusan yang sama. Dengan demikian empat uji mempunyai kepekaan yang sama untuk menolak H_0 ketika H_0 salah.

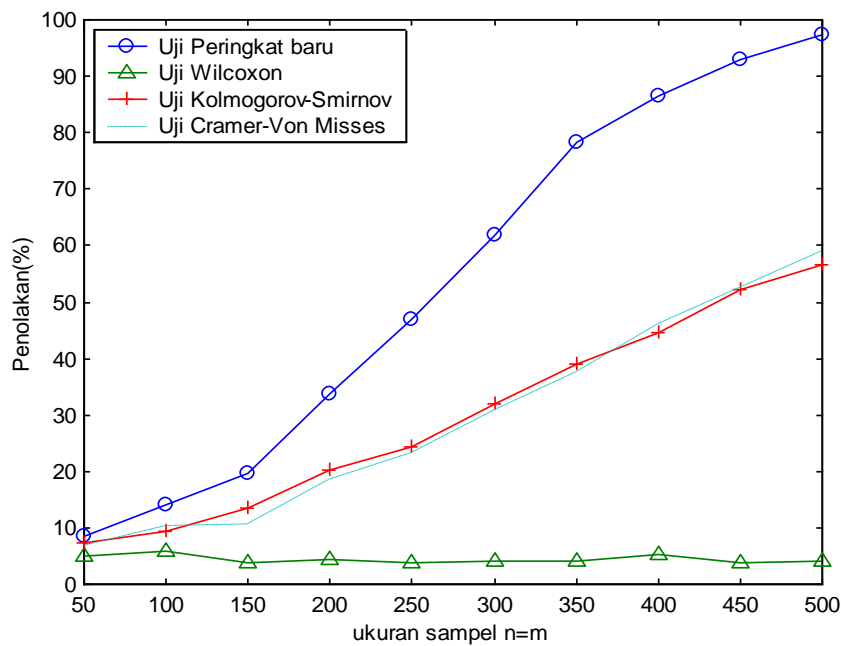
Bila diambil perbandingan nilai parameter tidak sama atau membesar, hasil simulasi memberikan keputusan yang tidak berbeda secara nyata, dengan kata lain keempat uji masih mempunyai kepekaan yang sama untuk menolak H_0 ketika H_0 salah. Dengan demikian dapat ditarik kesimpulan bahwa sampel yang mempunyai kemiripan karakteristik seperti distribusi eksponensial, empat uji tersebut memberikan keputusan yang sama.

Hasil simulasi keempat dapat dilihat pada gambar 4. Simulasi ini berasal dari sampel yang dibandingkan dari distribusi normal $N(0,1)$ dan distribusi seragam/uniform $UNIF(-\frac{1}{2}, \frac{1}{2})$. Pengambilan dua sampel berasal dari populasi yang berbeda, namun mean dan variansi dari dua sampel sama merupakan suatu kejadian yang berbeda dengan kejadian yang ada pada ketiga simulasi sebelumnya. Kedua sampel juga merupakan suatu distribusi yang mempunyai bentuk simetris.

Bila diperhatikan karakteristiknya secara sepintas tampak dua sampel tersebut berasal dari populasi yang sama. Namun demikian dugaan pernyataan tersebut harus dibuktikan secara statistik.



Gambar 3



Gambar 4

Ketika ukuran sampel diambil kecil, dalam hal ini $n=5$, tampak bahwa kedua sampel melalui empat uji menghasilkan keputusan yang sama. Bila diambil besarnya sampel 100, tampak dari gambar bahwa uji peringkat baru paling besar untuk menolak H_0 bila ternyata H_0 salah. Sedangkan uji Wilcoxon paling kecil untuk menolak H_0 ketika H_0 salah. Semakin besar sampel yang diambil tampak bahwa uji peringkat baru semakin kuat untuk menolak H_0 ketika H_0 salah.

Membesarnya ukuran sampel untuk uji Wilcoxon tidak meningkatkan terhadap penolakan H_0 ketika H_0 salah. Ini berarti uji Wilcoxon tidak mampu untuk membedakan dua yang sampel yang berbeda. Sedangkan untuk uji Cramer von Mises dan uji Kolmogorov-Smirnov tidak dapat dibandingkan untuk berbagai ukuran sampel sebab kedua grafik berpotongan secara bergantian, akan tetapi kedua uji ini memberikan keputusan yang sama.

KESIMPULAN

Berdasarkan pembahasan di atas dapat ditarik kesimpulan bahwa uji peringkat baru mempunyai kepekaan tertinggi untuk menolak H_0 ketika H_0 salah, ini berarti uji peringkat baru mempunyai kuasa uji paling baik diantara uji Cramer von Mises, uji Kolmogorov-Smirnov, dan uji Wilcoxon.

DAFTAR PUSTAKA

- Anderson, T.W. and Darling, D.A. 1952. *Asymptotic theory of certain goodness of fit criteria based on stochastic processes*. Annals of Mathematical Statistics. **23**, 193-212.
- Baumgartner, W., Wei β , P. and Schindler, H. 1998. *A new rank test for the general two-sample problem*. Biometric. **75**, 1129-1135.
- Conover, W. J. 1990. *Practical nonparametric Statistics*. Second Edition. John Wiley and Sons, New York.
- Efron, B. 1990. *More efficient bootstrap computations*. J. Amer. Statist. Assoc. **85**, 79-89.
- Ross, G.J and Adams, N.M. 2012. *Two Nonparametric Control Charts for detecting Arbitrary Distribution Changes*. Journal of Quality Technology, Vol. 44, N0.2.
- Shabri, A and Jemain, A.A. 2008. *Pengujian Statistik Anderson Darling bagi Taburan Nilai Ekstrim Teritlak*. Jurnal Matematika, Vol.24, N0.1, pp: 85-97.
- Steele, Michael. Chaceling, Janet. 2006. *Power of Discrete Goodness-of-Fit Test statistics For a Uniform Null Against a Selection of Alternative Distributions*. Journal of Simulation and Computation, Vol.35, pp:1067-1075.
- Sugiyanto dan Rini. I 2004. *Perbandingan uji Cramer-von Mises dan uji peringkat baru untuk dua sampel bebas melalui metode simulasi*. Majalah Math Info.
- Sugiyanto dan Ngapiningsih 2004. *Perbandingan uji Kolmogorov-Smirnov dan uji peringkat baru untuk dua sampel bebas melalui metode simulasi..* Majalah Math Info.
- Sugiyanto dan Wulan.S 2005. *Perbandingan uji Wilcoxon dan uji peringkat baru untuk dua sampel bebas melalui metode simulasi*. Majalah Math Info.
- Sugiyanto dan Hasih..P. 2006. *Perbandingan uji nonparametrik untuk dua sampel bebas melalui metode simulasi*, Laporan Penelitian Dosen Muda.
- Sugiyanto dan Isnandar S. 2007. *Perbandingan uji-uji untuk dua sampel bebas*, Laporan Penelitian Dosen Muda.
- Zhang, J. 2001. *Powerful Goodness-of-Fit and Multi-Sampel Test*. PhD Thesis. New-york Univercity. Toronto.