



Perbandingan Power of Test dari Uji Normalitas Metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson-Darling

Dyah Setyo Rini, Fachri Faisal

Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu, Indonesia

Diterima 13 April 2015 ; Disetujui 13 Juni 2015

Abstrak - Pelanggaran terhadap asumsi kenormalan dalam suatu prosedur statistika akan memberikan kesimpulan hasil yang diperoleh menjadi tidak valid. Tulisan ini akan membahas tentang perbandingan uji normalitas metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson-Darling berdasarkan nilai *power of test* dari masing-masing uji. *Power of test* dari masing-masing uji diperoleh dari *power* empiris melalui simulasi data. Data simulasi dibangkitkan dari distribusi t-student, gamma, chi-square, dan beta. Ukuran sampel data yang digunakan dalam simulasi adalah 10, 20, 30, 40, dan 50. Berdasarkan simulasi data, uji normalitas metode Bayesian MT_n lebih *powerful* apabila dibandingkan dengan uji Shapiro-Wilk (W), uji Cramer-von Mises (W_n^2), dan uji Anderson-Darling (A_n^2) untuk data yang berdistribusi t-student dengan derajat bebas 5 dan 10, Gamma (5), chi-square derajat bebas 20. Hal tersebut juga berlaku untuk sampel berukuran kecil. Performa dari uji normalitas metode Bayesian lebih lemah dari pada uji Shapiro-Wilk dalam menguji normalitas yang berasal dari sampel berdistribusi Beta, $B(a,b)$, dengan $a=0,8$ dan $1,5$; $b=0,8$ dan $0,9$.

Kata Kunci: *power empiris, simulasi data, uji normalitas*

1. Pendahuluan

Sebagian besar penelitian dalam bidang statistika berhubungan dengan pengujian asumsi distribusi, baik secara teori maupun praktik di lapangan. Salah satu pengujian asumsi distribusi yang paling sering digunakan adalah pengujian asumsi kenormalan, artinya menguji apakah sampel data yang digunakan memiliki distribusi normal atau tidak. Distribusi normal merupakan distribusi yang paling penting dalam statistika untuk berbagai alasan. Analisis data statistika secara parametrik sangat memerlukan asumsi kenormalan. Diantaranya, data sering diaproksimasi ke dalam bentuk kurva normal dan beberapa distribusi akan menjadi distribusi normal secara asimtotik. Beberapa statistik uji, seperti uji t, uji F untuk kehomogenan varian, uji koefisien regresi, dan analisis varian menggunakan asumsi bahwa sampel data yang digunakan haruslah memiliki distribusi normal. Pelanggaran terhadap asumsi kenormalan dalam suatu prosedur statistika akan memberikan kesimpulan hasil yang diperoleh menjadi tidak valid.

Banyak literatur yang telah membahas tentang uji-uji normalitas. Referensi [1] merupakan salah satu literatur yang membahas tentang berbagai uji normalitas dan prosedurnya. Referensi [2] mengkaji tentang uji normalitas yang diperoleh berdasarkan metode Bayesian, yaitu dengan melihat perbedaan atau ketidaksesuaian antara sampel pengamatan dan sampel prediksi yang berasal dari distribusi posterior prediktif. Referensi [3] membahas perbandingan *power* dari uji Kolmogov-Smirnov, uji Shapiro-Wilk, uji Liliefors, dan uji Anderson-Darling. Menurut [4] uji Anderson-Darling dan uji Shapiro-Wilk merupakan uji normalitas yang paling *powerful*. Setiap uji normalitas terkadang memberikan kesimpulan yang berbeda. Oleh karena itu, masing-masing uji tersebut memiliki asumsi atau kondisi tertentu dari sampel data agar dapat memperoleh hasil yang diharapkan. Berdasarkan hal tersebut, penulis akan membahas perbandingan *power of test* pada uji normalitas metode Bayesian, uji Shapiro-Wilk, uji Cramer-von Mises, dan uji Anderson-Darling berdasarkan nilai *power* empiris. Penulis berharap tulisan ini dapat memberikan tambahan wawasan pengetahuan dan dapat memberikan gambaran tentang uji normalitas metode Bayesian kepada para pembaca.

Uji Anderson-Darling

Uji Anderson-Darling untuk *goodness-of-fit* digunakan untuk mengetahui distribusi dari suatu sampel. Statistik uji Anderson-Darling didefinisikan sebagai berikut [1]:

$$A^2 = -n - n^{-1} \sum_{i=1}^n [2i - 1] \left[\log(p_{(i)}) + \log(1 - p_{(n-i+1)}) \right]$$

dengan $p_{(i)} = F^0(x_{(j)})$.

$F^0(x_{(j)})$ merupakan distribusi hipotesis dari sampel terurut

$$x_{(1)} < x_{(2)} < \dots < x_{(n)}$$

Uji Shapiro-Wilk

Uji Shapiro-Wilk untuk normalitas ini dikembangkan oleh Samuel Shapiro dan Martin Wilk pada tahun 1965. Pada saat ini, uji Shapiro-Wilk menjadi uji normalitas yang lebih disukai karena memiliki kekuatan uji yang lebih baik dibandingkan uji-uji alternatif dari bermacam-macam *range*. Uji ini tergantung pada korelasi antara data yang diberikan dan kecocokan angka normalnya.

Statistik uji Shapiro-Wilk dirumuskan sebagai berikut [1]:

$$W = \frac{b^2}{(n-1)s^2}$$

dengan

$$b^2 = \sum_{i=1}^{\frac{n}{2}} a_{n-i+1} (x_{(n-i+1)} - x_{(i)}),$$

$x_{(i)}$ merupakan nilai sampel terbesar ke- i dari sampel terurut

$$x_{(1)} < x_{(2)} < \dots < x_{(n)},$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Uji Cramer-von Mises

Uji Cramer-von Mises dikembangkan oleh Cramer (1928), von Mises (1931), dan Smirnov (1936). Statistik uji dari uji Cramer-von Mises adalah sebagai berikut:

$$W_n^2 = \frac{1}{12n} + \sum_i \left(P_{(i)} - \frac{2i-1}{2n} \right)^2$$

Uji Normalitas Metode Bayesian

Pendekatan Bayesian dalam statistik secara fundamental, berbeda dengan pendekatan klasik. Dalam pendekatan klasik, parameter θ adalah besaran tetap yang tidak diketahui. Sampel random X_1, X_2, \dots, X_n diambil dari populasi berindeks θ dan berdasarkan nilai-nilai terobservasi dalam sampel, didapat pengetahuan tentang θ .

Dalam pendekatan Bayesian, θ dipandang sebagai besaran yang variasinya digambarkan dengan distribusi probabilitas (disebut distribusi prior) yang bersifat subyektif, yakni berdasarkan pada keyakinan seseorang dan dirumuskan sebelum data diambil. Kemudian sampel diambil dari populasi berindeks θ dan distribusi prior disesuaikan dengan informasi sampel ini. Prior yang telah disesuaikan disebut distribusi posterior [5].

Menurut [6], Bila data x telah diamati, maka dapat diprediksi suatu pengamatan y melalui proses yang sama. Distribusi dari y disebut dengan distribusi posterior prediktif, yakni:

$$h(y | x) = \int f(y | \theta) \pi(\theta | x) d\theta$$

Misalkan E_1, E_2, \dots, E_n merupakan sampel yang diperoleh dari distribusi posterior prediktif $h(\square \mathbf{x})$, maka (E_1, E_2, \dots, E_n) dapat dianggap sebagai suatu sampel prediksi dari sampel (X_1, X_2, \dots, X_n) .

Statistik uji untuk uji normalitas metode Bayesian adalah:

$$MT_n := MT_n(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n \frac{E(X_{(i)} - E_{(i)})^2}{S^2} \quad (1)$$

$$MU_n := MU_n(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n E \left(\frac{X_{(i)} - E_{(i)}}{\sigma} \right)^2 \quad (2)$$

Nilai ekspektasi dalam persamaan (1) diperoleh dari distribusi (E_1, E_2, \dots, E_n) apabila (X_1, X_2, \dots, X_n) diketahui. Sedangkan nilai ekspektasi pada persamaan (2) diperoleh dari distribusi posterior bersama dari (E_1, E_2, \dots, E_n) dan σ^2 bersyarat (X_1, X_2, \dots, X_n) [2].

2. Metode Penelitian

Penelitian ini menggunakan studi literatur. Simulasi data digunakan dalam menghitung power empiris dari uji normalitas metode Bayesian, uji Shapiro-Wilk, uji Cramer-von Mises, dan uji Anderson-Darling dengan bantuan program R.

3. Hasil dan Pembahasan

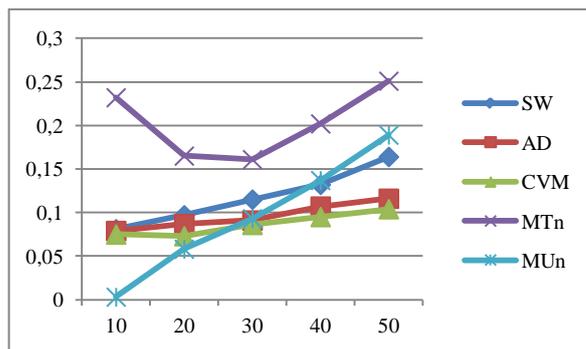
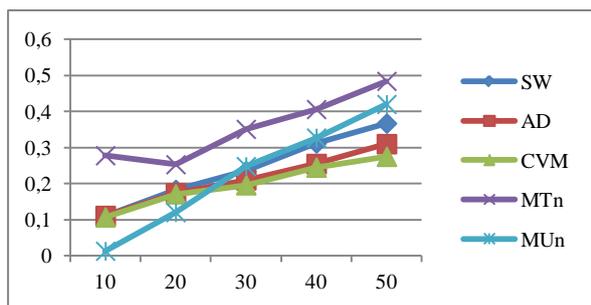
Perbandingan power of test pada uji normalitas metode Bayesian (MT_n dan MU_n), uji Shapiro-Wilk (W), uji Cramer-von Mises (W_n^2), dan uji Anderson-Darling (A_n^2) dapat dilihat berdasarkan nilai power empiris dari masing-masing uji. Power empiris diperoleh melalui simulasi data dengan bantuan program R. Simulasi data dilakukan dalam 1000 kali perulangan. Taraf signifikansi yang digunakan adalah 5%. Serta ukuran sampel yang digunakan untuk simulasi data adalah $n = 10, 20, 30, 40, 50$.

Nilai Power Empiris dari masing-masing Uji Normalitas dengan Menggunakan Sampel Berdistribusi t-student.

Tabel 1 menampilkan nilai power empiris dari uji normalitas metode Bayesian, uji Shapiro-Wilk, uji Cramer-von Mises, dan uji Anderson-Darling dengan menggunakan sampel berdistribusi t-student. Derajat bebas yang digunakan adalah 5 dan 10, serta nilai kurtosisnya masing-masing adalah 7,32 dan 4,04.

Tabel 1. Nilai Power Empiris dari Uji Normalitas Metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson Darling dengan menggunakan sampel distribusi t-student

Distribusi Alternatif	Kurtosis	Ukuran Sampel (n)	Power Empiris				
			W	A_n^2	W_n^2	MT_n	MU_n
t(5)	7.32	10	0.111	0.111	0.107	0.278	0.013
		20	0.183	0.174	0.171	0.253	0.120
		30	0.235	0.210	0.196	0.351	0.248
		40	0.312	0.255	0.245	0.406	0.327
		50	0.367	0.310	0.275	0.484	0.420
t(10)	4.04	10	0.081	0.079	0.075	0.232	0.003
		20	0.097	0.087	0.073	0.165	0.058
		30	0.115	0.091	0.086	0.161	0.093
		40	0.132	0.107	0.095	0.202	0.137
		50	0.164	0.116	0.104	0.251	0.189



Gambar 1. Power Empiris dari Uji Normalitas Metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson Darling dengan menggunakan sampel berdistribusi t-student

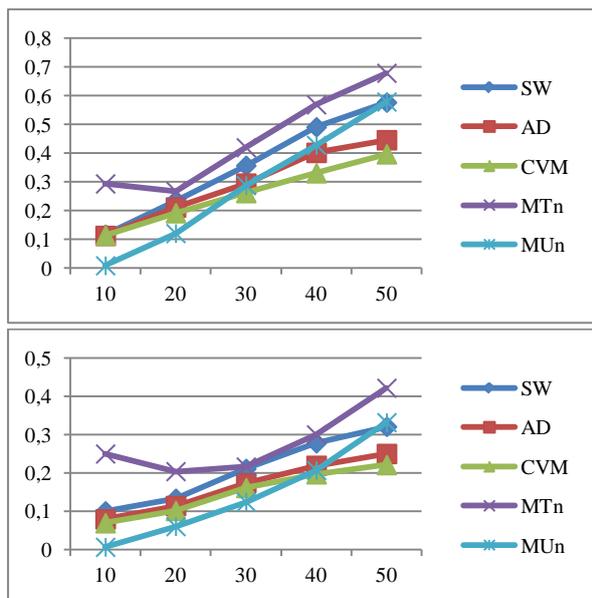
Berdasarkan Tabel 1 terlihat bahwa nilai power empiris dari MT_n tidak beraturan untuk ukuran sampel 10, 20, dan 30. Hal tersebut dipengaruhi oleh perbedaan derajat bebas dari distribusi t-student yang digunakan dalam simulasi. Akan tetapi, statistik uji MT_n memiliki nilai power empiris yang paling besar dibandingkan dengan W, A_n^2, W_n^2 untuk $n = 10, 20, 30, 40, 50$. Sedangkan uji normalitas MU_n terlihat lebih powerful dibandingkan dengan W, A_n^2, W_n^2 untuk ukuran sampel yang lebih besar dari 30.

Nilai Power Empiris dari masing-masing Uji Normalitas dengan Menggunakan Sampel Berdistribusi Gamma dan Chi-Square.

Pada Tabel 2 terlihat nilai power empiris dari masing-masing uji normalitas dengan menggunakan sampel yang berdistribusi Gamma(5) dengan nilai skewness dan kurtosisnya adalah 0,90 dan 4,21 serta sampel yang berasal dari distribusi Chi-Square dengan derajat bebas 20 dan nilai skewness dan kurtosisnya adalah 0,64 dan 3,59.

Tabel 2. Nilai Power Empiris dari Uji Normalitas Metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson Darling dengan menggunakan sampel distribusi Gamma dan Chi-Square

Distribusi Alternatif	Skewness	Kurtosis	Ukuran Sampel (n)	Power Empiris				
				W	A_n^2	W_n^2	MT_n	MU_n
Gam(5)	0.90	4.21	10	0.117	0.113	0.114	0.294	0.009
			20	0.235	0.212	0.192	0.268	0.121
			30	0.357	0.295	0.262	0.420	0.288
			40	0.491	0.402	0.331	0.569	0.428
			50	0.577	0.446	0.397	0.679	0.579
$\chi_{(20)}^2$	0.64	3.59	10	0.100	0.079	0.069	0.250	0.006
			20	0.133	0.114	0.101	0.204	0.060
			30	0.211	0.174	0.161	0.216	0.124
			40	0.278	0.219	0.197	0.300	0.207
			50	0.321	0.250	0.221	0.422	0.332



Gambar 2. Power Empiris dari Uji Normalitas Metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson Darling dengan menggunakan sampel berdistribusi Gamma dan Chi-Square

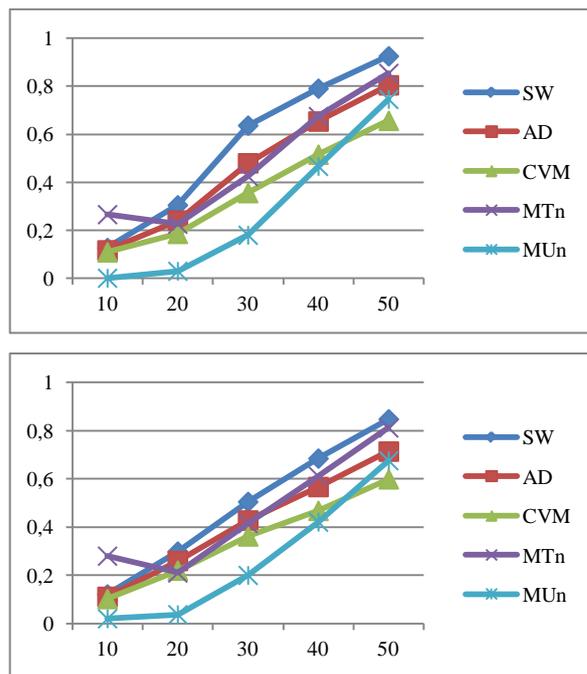
Tabel 2 menunjukkan bahwa uji normalitas dengan menggunakan metode Bayesian MU_n lebih *powerful* dibandingkan dengan uji normalitas lainnya untuk ukuran sampel 50. Sedangkan secara keseluruhan uji normalitas MT_n lebih *powerful* apabila dibandingkan dengan statistik uji lainnya untuk sampel berdistribusi Gamma dan Chi-Square pada taraf signifikansi 5%.

Nilai Power Empiris dari masing-masing Uji Normalitas dengan Menggunakan Sampel Berdistribusi Beta.

Tabel 3 menampilkan nilai power empiris dari masing-masing uji normalitas dengan menggunakan sampel berdistribusi Beta, $B(a,b)$, dengan $a=0,8$ dan $1,5$; $b=0,8$ dan $0,9$. Nilai skewness yang digunakan adalah 0 dan $-0,43$ serta nilai kurtosisnya adalah $1,69$ dan $2,11$.

Tabel 3. Nilai Power Empiris dari Uji Normalitas Metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson Darling dengan menggunakan sampel berdistribusiBeta

Distribusi Alternatif	Skewness	Kurtosis	Ukuran Sampel (n)	Power Empiris				
				W	A_n^2	W_n^2	MT_n	MU_n
$B(0.8, 0.8)$	0	1.69	10	0.127	0.117	0.109	0.266	0.001
			20	0.305	0.241	0.186	0.226	0.030
			30	0.637	0.480	0.356	0.425	0.180
			40	0.791	0.654	0.516	0.675	0.467
			50	0.926	0.805	0.657	0.855	0.746
$B(1.5, 0.9)$	-0.43	2.11	10	0.121	0.111	0.102	0.280	0.020
			20	0.299	0.260	0.220	0.210	0.037
			30	0.505	0.428	0.362	0.415	0.200
			40	0.685	0.567	0.468	0.611	0.421
			50	0.848	0.714	0.599	0.812	0.676



Gambar 3. Power Empiris dari Uji Normalitas Metode Bayesian, Uji Shapiro-Wilk, Uji Cramer-von Mises, dan Uji Anderson Darling dengan menggunakan sampel berdistribusi Beta

Hasil yang disajikan menunjukkan bahwa uji Shapiro-Wilk lebih *powerful* dibandingkan dengan metode Bayesian. Hal itu dapat dilihat dari nilai *power* empiris dari uji-uji tersebut yang lebih besar dari pada metode Bayesian. Dengan kata lain, metode Bayesian lebih lemah dari pada metode klasik terutama uji Shapiro-Wilk dalam menguji normalitas dengan sampel berdistribusi Beta pada taraf signifikansi 5%.

4. Kesimpulan

Berdasarkan uraian dan perhitungan yang dilakukan, dapat disimpulkan bahwa uji normalitas metode Bayesian MT_n lebih *powerful* apabila dibandingkan dengan uji Shapiro-Wilk (W), uji Cramer-von Mises (W_n^2), dan uji Anderson-Darling (A_n^2). Hal tersebut juga berlaku untuk sampel berukuran kecil. Performa dari uji normalitas metode Bayesian lebih lemah dari pada uji Shapiro-Wilk dalam menguji normalitas yang berasal dari sampel berdistribusi Beta.

Daftar Pustaka

[1] Thode, H.C. , 2002, Testing for Normality, New York: Marcel Dekker.
 [2] He, D.J. dan Xu, X.Z., 2013, "A Goodness-of-Fit Testing Approach for Normality Based On The

Posterior Predictive Distribution”, *Test*, vol. 22, Issue, pp. 1-18.

- [3] Razali, N.M. dan Wah, Y.B., 2011 , “Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Liliefors and Anderson-Darling Tests”, *Journal of Statistical Modeling and Analytics*, vol. 2, pp. 21-33.
- [4] Stephens, M.A., 1974 , “EDF Statistics for Goodness of Fit and Some Comparisons”, *Journal of American Statistical Association*, vol. 69, pp. 730-737.
- [5] Subanar, 2006, *Inferensi Bayesian*, Universitas Terbuka, Jakarta.
- [6] Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004, *Bayesian Data Analysis 2nd Edition*, Chapman and Hall, New York.