

STUDI SIMULASI : PERBANDINGAN UJI WELCH DAN UJI COHRAN-COX PADA MASALAH BEHRENS-FISHER

Sudartianto

Departemen Statistika, FMIPA UNPAD
sudartianto@unpad.ac.id

Yudhi Andriyana

Departemen Statistika, FMIPA UNPAD
y.andriyana@unpad.ac.id

Riaman

Departemen Matematika, FMIPA UNPAD
riaman@unpad.ac.id

ABSTRACT. *Behrens-Fisher problem has a lot of test, two of them are Cochran-Cox test and Welch test. This study will examine the stability of both tests by comparing the significance levels of both tests to a specific (given) significance level. The result shows that the Welch test is more stable than the Cochran-Cox test which means that theoretically Welch test is better than Cochran-Cox test.*

Keywords: *Behrens-Fisher problem, Cochran-Cox test, Welch test.*

ABSTRAK. Masalah Behrens-Fisher mempunyai banyak sekali uji, dua di antaranya adalah uji Cochran-Cox dan uji Welch. Penelitian ini akan mengkaji kedua uji tersebut, dengan melihat kestabilan nilai taraf signifikansi, dari kedua uji tersebut, kemudian dibandingkan dengan taraf signifikansi nominal yang kita kehendaki, untuk melihat mana yang mempunyai taraf signifikansi yang lebih stabil. Hasilnya terlihat bahwa uji welch lebih stabil dibandingkan dengan uji Cochran-Cox, itu berarti secara teori uji Welch lebih baik dari pada uji Cochran-Cox.

Kata Kunci: Masalah Behren-Fisher, Uji Cochran-Cox, Uji Welch.

1. PENDAHULUAN

Salah satu masalah yang sering timbul bagi para peneliti yang menggunakan statistika, adalah membandingkan rata-rata dari dua populasi. Ketika membandingkan rata-rata dari dua populasi kadang-kadang timbul persoalan, yaitu jika varians dari kedua populasi tersebut tidak sama, masalah ini sering disebut sebagai masalah Behrens-Fisher (Behren,1964) dan (Fisher, 1941), selain itu dalam menguji kesamaan dua rata-rata juga ada asumsi bahwa kedua populasi tersebut mengikuti distribusi normal, sedangkan kedua sampel acak yang diambil dari masing-masing populasi adalah bebas.

Karena banyaknya pengujian untuk masalah Behrens-Fisher ini, maka kita perlu mencari yang manakah dari pengujian-pengujian tersebut yang merupakan pengujian terbaik, caranya dengan melihat kestabilan dari ukuran daerah kritis untuk masing-masing pengujian, yang terbaik adalah pengujian yang mempunyai ukuran daerah kritis yang lebih stabil (Lee A.F.S. and Gurland J. 1975). Selanjutnya juga dari ukuran-ukuran daerah kritis yang stabil kita akan bandingkan kuasa uji dari pengujian, untuk lebih meyakinkan kita mana yang merupakan pengujian terbaik, di dalam makalah ini kami ingin membandingkan dua uji dalam masalah Behrens-Fisher yang banyak dipakai karena kesederhanaan dan kemudahannya dalam pemakaian, yaitu uji Cochran-Cox dan uji Welch.

Selanjutnya, tujuan dari penelitian ini yang ingin dicapai yaitu ingin diketahui dari kedua metode tersebut, metode manakah yang merupakan metode terbaik untuk masalah Behrens-Fisher.

2. METODE UNTUK MENGUJI KESAMAAN DUA RATA-RATA

Misalkan kita mempunyai dua populasi yang masing-masing berdistribusi normal dengan rata-rata μ_1 dan μ_2 serta variansnya σ_1^2 dan σ_2^2 . Dari tiap populasi diambil sampel acak yang saling bebas, misalkan sampel acak yang diambil dari populasi kesatu berukuran n_1 , dengan rata-rata \bar{X}_1 dan varians S_1^2 , sedangkan yang diambil dari populasi kedua berukuran n_2 dengan rata-rata \bar{X}_2 dan varians S_2^2 . Berdasarkan sampel tersebut kita akan menguji kesamaan rata-rata dari dua populasi. Pada umumnya σ_1^2 dan σ_2^2 tidak diketahui, oleh sebab itu harganya ditaksir oleh S_1^2 dan S_2^2 . Untuk menguji kesamaan dua rata-rata dari dua populasi normal dengan varians tidak sama digunakan statistik uji yaitu :

$$t' = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^{1/2}} \quad (1)$$

Statistik uji ini merupakan statistik uji pendekatan sebagai solusi dari masalah Behrens-Fisher karena tidak ada yang eksak. Yang menjadi persoalan adalah bagaimana menentukan distribusi dari t' seperti pada Persamaan (1). Untuk memudahkan pembahasan selanjutnya digunakan notasi sebagai berikut :

n_i = Ukuran sampel dari populasi ke i ($i=1,2$).

\bar{X}_i = Rata-rata sampel dari populasi ke i.

$d = \bar{X}_1 - \bar{X}_2$ = Selisih rata-rata sampel dari populasi ke 1 dengan rata-rata sampel dari populasi ke 2.

S_i^2 = varians sampel dari populasi ke i

$f_i = n_i - 1$ = derajat kebebasan dari sampel ke i.

$S_d^2 = S_1^2/n_1 + S_2^2/n_2$ = Gabungan varians dari sampel.

$\sigma_d^2 = \sigma_1^2/n_1 + \sigma_2^2/n_2$ = Gabungan varians populasi.

$\delta = \mu_1 - \mu_2$ = Selisih antara rata-rata populasi ke-1 dengan rata-rata populasi ke 2.

Berdasarkan notasi-notasi di atas, maka Persamaan (1), dapat disederhanakan menjadi :

$$t' = (d - \delta) / sd \quad (2)$$

2.1 Uji Cochran-Cox

Cara yang sering digunakan untuk menguji hipotesis tersebut di atas dikemukakan oleh Cochran dan Cox, dalam (Hogg and Craig, 1978) yang penjelasannya adalah sebagai berikut : titik kritisnya merupakan rata-rata dengan bobot untuk $t_{(1-\alpha/2; f_1)}$ dan $t_{(1-\alpha/2; f_2)}$ dalam hal ini bobotnya adalah $\lambda_1 S_1^2$ dan $\lambda_2 S_2^2$, untuk $\lambda_i = 1/n_i$; $i = 1, 2$ yakni :

$$V = \frac{(\lambda_1 S_1^2) t_{(1-\frac{\alpha}{2}; f_1)} + (\lambda_2 S_2^2) t_{(1-\frac{\alpha}{2}; f_2)}}{(\lambda_1 S_1^2) + (\lambda_2 S_2^2)} \quad (3)$$

Yang dalam hal ini, $t_{(1-\alpha/2; f_1)}$ dan $t_{(1-\alpha/2; f_2)}$ masing-masing merupakan nilai yang diperoleh dari tabel distribusi student t dengan derajat kebebasan $f_1 = n_1 - 1$ dan $f_2 = n_2 - 1$ dengan peluang $1-\alpha/2$. Dengan demikian, kriteria pengujiannya adalah : Terima H_0 jika $-V \leq t' \leq V$, tolak H_0 dalam keadaan yang lainnya (Lauer G. N., 1971).

2.2 Uji Welch

Jika $X_{11}, X_{12}, \dots, X_{1n_1}$ dan $X_{21}, X_{22}, \dots, X_{2n_2}$ merupakan dua sampel acak yang saling bebas dari masing-masing populasi yang berdistribusi normal dengan rata-rata μ_1 dan μ_2 serta variansnya σ_1^2 dan σ_2^2 , maka berdasarkan uji Welch

dalam (Welch, B.L. 1949) kriteria uji untuk uji Welch adalah terima H_0 jika $-t(1-\alpha/2; \hat{f}) \leq t' \leq t(1-\alpha/2; \hat{f})$ dan tolak H_0 dalam keadaan yang lainnya, untuk

$$\hat{f} = \frac{(S_1^2/n_1 + S_2^2/n_2)}{f_1^{-1}(S_1^2/n_1)^2 + f_2^{-1}(S_2^2/n_2)^2} \quad (4)$$

Uji Welch ini menurut (Derrick B., Toher D., and White P. 2016) merupakan uji yang robust.

3. CARA MENCARI UKURAN DAERAH KRITIS DAN KUASA UJI

Bila $R = \sigma_1^2 / \sigma_2^2$ diketahui ukuran daerah kritis dari uji yang tergantung pada R yaitu $S(R)$ dapat dicari (Lauer G.N., 1971), untuk kemudahan kita gunakan uji pihak kanan,

$$\begin{aligned} S(R) &= P[\text{Tolak } H_0 \mid \delta=0, R] \\ &= P[V^2 > [\frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}]^2 \mid \delta = 0, R] \end{aligned}$$

dengan $w_1 = S_1^2/n_1$; $w_2 = S_2^2/n_2$; $t_1 = t_{(1-\frac{\alpha}{2}; f_1)}$ dan $t_2 = t_{(1-\frac{\alpha}{2}; f_2)}$. Bila R diketahui,

$S(R)$ dapat dicari. Sudah diketahui bahwa :

$$\begin{aligned} \frac{S_i^2 \cdot f_i}{\sigma_i^2} &\sim \chi^2_{(f_i)} ; i = 1, 2 \\ \frac{(\bar{x}_1 - \bar{x}_2)^2}{\sigma_1^2/n_1 + \sigma_2^2/n_2} &\sim \chi^2_{(1)} \end{aligned}$$

Semua χ^2 di atas saling bebas

$$S(R) = P\left[\frac{C \chi^2_{(1)}}{v_1 x^2_{(f_1)} + v_2 x^2_{(f_2)}} > \emptyset^2(x^2_{(f_1)}, x^2_{(f_2)}, \alpha)\right]$$

dengan

$$\begin{aligned} c &= \frac{R}{n_1} + \frac{1}{n_2}; v_1 = \frac{R}{n_1 f_1}; v_2 = \frac{1}{n_2 f_2} \\ \emptyset^2(x^2_{(f_1)}, x^2_{(f_2)}, \alpha) &= \left[\frac{v_1 t_1 x^2_{(f_1)} + v_2 t_2 x^2_{(f_2)}}{v_1 x^2_{(f_1)} + v_2 x^2_{(f_2)}} \right]^2 \end{aligned}$$

Sehingga bila diketahui α, n_1, n_2 dan R maka ukuran daerah kritis dari uji Cochran-Cox dapat dicari, sedangkan power dari suatu uji dapat ditulis sebagai suatu fungsi dari R , dengan pola yang sama dengan ukuran daerah kritis, ditunjukkan oleh $P(R, \lambda)$ terlihat sebagai bentuk

$$P(R, \lambda) = P[\text{Tolak } H_0 \mid \delta] \\ = P\left[\frac{C \chi'^2(1, \lambda)}{v_1 x_{(f_1)}^2 + v_2 x_{(f_2)}^2} > \phi^2(x_{(f_1)}^2, x_{(f_2)}^2, \alpha)\right]$$

Dimana $\chi'^2(1, \lambda)$ menggambarkan peubah acak χ^2 non central dengan $dk = 1$ dan parameter noncentralitas λ , mudah dibuktikan bahwa ;

$$\lambda = \frac{\delta^2}{\sigma_1^2/n_1 + \sigma_2^2/n_2} = \left(\frac{\delta}{\sigma_2}\right)^2 \left(\frac{n_1 n_2}{R n_2 + n_1}\right)$$

Sedangkan untuk uji Welch, untuk mencari ukuran daerah kritis dan kuasa ujinya hanya mencari nilai peluang dari : $P[\text{Tolak } H_0 \mid \delta] = P(t' > t(1 - \alpha/2; \hat{f}))$ dengan cara menggunakan interpolasi linier apabila diketahui nilai-nilai n_1 , n_2 , α dan R .

4. PERBANDINGAN UJI COCHRAN-COX DAN UJI WELCH

Sesuai dengan teori statistika diketahui bahwa pengujian terbaik adalah pengujian yang mempunyai kuasa uji yang terbesar untuk ukuran daerah kritis atau taraf signifikansi α hitung yang sama. Berdasarkan kriteria pengujian terbaik di atas, ternyata dalam hal ini kita memperoleh dua pengujian dengan nilai α hitung yang berbeda oleh sebab itu diambil perubahan kriteria yaitu kita mencari nilai α hitung yang stabil untuk berbagai nilai n_1 , n_2 , α dan R (Lee, A.F.S. and Gurland J. 1975) sehingga diperoleh hasil seperti Tabel 1. berikut :

Tabel 1. Perbandingan Ukuran Daerah Kritis antara Uji Cochran-Cox dan Uji Welch

n_1	n_2	$\alpha = 0.01 ; R = 1$		$\alpha = 0.01 ; R = 4$	
		Uji Cochran-Cox	Uji Welch	Uji Cochran-Cox	Uji Welch
3	3	0,0009	0,0009	0,0009	0,0009
3	5	0,0015	0,0016	0,0037	0,0038
3	7	0,0029	0,0030	0,0065	0,0066
3	9	0,0042	0,0042	0,0083	0,0083
5	3	0,0015	0,0016	0,0012	0,0014
5	5	0,0018	0,0019	0,0028	0,0030
5	7	0,0025	0,0026	0,0046	0,0047
5	9	0,0033	0,0035	0,0060	0,0062

Lanjutan Tabel 1.

n ₁	n ₂	$\alpha = 0.05 ; R = 1$		$\alpha = 0.05 ; R = 4$	
		Uji Cochran-Cox	Uji Welch	Uji Cochran-Cox	Uji Welch
3	3	0,0126	0,0127	0,0171	0,0171
3	5	0,0204	0,0205	0,0317	0,0318
3	7	0,0265	0,0265	0,0388	0,0389
3	9	0,0310	0,0311	0,0425	0,0427
5	3	0,0204	0,0205	0,0190	0,0191
5	5	0,0241	0,0242	0,0301	0,0302
5	7	0,0281	0,0283	0,0364	0,0365
5	9	0,0312	0,0314	0,0401	0,0402

n ₁	n ₂	$\alpha = 0.10 ; R = 1$		$\alpha = 0.10 ; R = 4$	
		Uji Cochran-Cox	Uji Welch	Uji Cochran-Cox	Uji Welch
3	3	0,0432	0,0432	0,0537	0,0539
3	5	0,0577	0,0579	0,0756	0,0757
3	7	0,0668	0,0670	0,0844	0,0845
3	9	0,0728	0,0729	0,0888	0,0889
5	3	0,0577	0,0579	0,0566	0,0567
5	5	0,0656	0,0658	0,0750	0,0752
5	7	0,0716	0,0718	0,0832	0,0834
5	9	0,0758	0,0760	0,0876	0,0879

Dari hasil di atas, tampak terlihat bahwa uji Welch memiliki nilai α hitung yang lebih stabil yang nilainya mendekati nilai α nominal sebesar 0.01, 0.05, 0.10 bila $R = 1$ dan $R=4$, untuk memperkuat kesimpulan ini, kita lihat bagaimana kuasa uji untuk kedua uji tersebut, hasilnya seperti Tabel 2 berikut ini.

Tabel 2. Perbandingan Kuasa Uji Antara Uji Cochran-Cox Dengan Uji Welch

n_1	n_2	λ	δ/σ_2	Uji Cochran-Cox	Uji Welch
3	3	3,5	1,53	0,105	0,109
		9,0	2,45	0,286	0,288
		24,0	4,00	0,695	0,698
3	5	4,2	1,50	0,201	0,199
		15,0	2,83	0,621	0,630
		30,0	4,00	0,877	0,879
3	7	4,8	1,51	0,254	0,256
		15,0	2,67	0,626	0,630
		3,50	4,08	0,906	0,917
3	9	5,1	1,51	0,277	0,278
		15,0	2,58	0,621	0,631
		35,0	3,94	0,899	0,899
5	3	9,5	1,15	0,123	0,126
		7,0	1,93	0,328	0,337
		20,0	3,27	0,739	0,747
5	5	2,6	1,02	0,180	0,189
		8,0	1,79	0,545	0,554
		20,0	2,83	0,937	0,947
5	7	2,8	0,98	0,219	0,228
		7,0	1,55	0,528	0,536
		20,0	2,62	0,954	0,962
5	9	3,0	0,97	0,247	0,255
		8,0	1,58	0,606	0,611
		90,0	2,49	0,955	0,960

Berdasarkan hasil tersebut di atas terlihat bahwa kuasa uji untuk uji Welch mempunyai nilai yang lebih besar dibandingkan dengan uji Cochran-Cox, dengan demikian dapat dikatakan secara teori uji Welch lebih baik dibandingkan dengan uji Cochran-Cox.

5. KESIMPULAN

Dari uraian yang telah diberikan untuk masalah Behrens-Fisher maka dengan ini dapat diambil beberapa kesimpulan.

- Masalah Behrens-Fisher pada umumnya dipergunakan jika kedua varians tersebut berbeda dan tidak diketahui. Sedangkan statistik uji untuk masalah Behrens-Fisher sampai sekarang merupakan statistik uji pendekatan.
- Untuk mencari pengujian terbaik dalam masalah Behrens -Fisher, maka kita perlu mencari pengujian yang mempunyai taraf signifikansi atau ukuran daerah kritis dari pengujian yang stabil dan kuasa ujinya cukup tinggi.

- c. Dari hasil di atas terlihat bahwa uji Welch lebih baik dari pada uji Cochran-Cox.

DAFTAR PUSTAKA

- Aspin, A. A., *An Examination and Further Development of a Formula Arising in the Problem of Comparing Two Mean Values*, *Biometrika*, **35** (1983), 88-96.
- Behrens, W. U., *The Comparison of Means of Independent Normal Distribution with Different Variances*, *Biometrics*, **20** (1964), 16-27.
- Derrick, B., Toher, D., dan White, P., *Why Welch's Test is Type I Error Robust*, *The Quantitative Methoda for Psychology*, 2016, 30-38.
- Ekiz, M. dan Gamgam H., *On The Comparison of The Welch Test and The Single-state Test : Simulation Study*, *Commun. Fac. Sci. Univ. Ank Series A1*. **56**(2) (2007), 51-61.
- Fisher, R. A., *The Asymptotic to Behrens Integral, with Further Tables for the d Test of Significance*, *Annal Eugenics*, **11**(2) (1941), 141-172.
- Hogg, R. V. dan Craig, A. T., *Introduction to Mathematical Statistics*, Fourth Edition, Macmillan Publishing Co. Inc., New York, , 1978.
- Lauer G. N., *Power of Cochran's Test in Behrens-Fisher Problems*, Dissertation, Iowa State University, 1971.
- Lee, A. F. S. dan Gurland J., *Size and Power of Tests for Equality of Means of Two Normal Population with Unequal Variances*, *Journal of- American Statistical Association*, **70** (1975), 933-941.
- Wang Y. Y., *Probabilities of the Type I Errors of the Welch Tests for the Behrens-Fisher Problem*, *Journal of American Statistical Association*, **66** (1971), 605-608.
- Welch, B. L., *The Generalization of 'Students' Problem When Several Different Population Variances Are Involved*, *Biometrika*, **36** (1949), 293-296.