Statistika, Vol. 8 No. 2, 119 – 127
Nopember 2008

Uji Permutasi untuk Masalah Dua Sampel Saling Bebas: Studi Kasus di LAFI-DITKES AD Bandung Jawa Barat

DANANG SETIAWAN DAN ACENG K. MUTAQIN

Program Studi Statistika Universitas Islam Bandung
Jl. Purnawarman No. 63 Bandung 40116
Email: shidiq_03@yahoo.com

ABSTRAK

Makalah ini membahas prosedur uji permutasi untuk masalah dua sampel saling bebas. Prosedur ini merupakan alternatif dari uji Wilcoxon-Mann-Whitney untuk menangani masalah dua sampel saling bebas jika datanya tidak berdistribusi normal. Data yang akan digunakan untuk mengaplikasikan metode tersebut diterapkan pada data sekunder mengenai waktu kegagalan dua merk mesin stripping yang digunakan dalam produksi obat-obatan di Lembaga Farmasi-Direktorat Kesehatan Angkatan Darat Bandung Jawa Barat selama periode Maret 2004 sampai Juli 2005. Berdasarkan hasil perhitungan menunjukan bahwa rata-rata waktu perawatan mesin stripping merk Chuan Yung sama dengan rata-rata waktu perawatan mesin stripping merk Ganson.

Kata Kunci: *uji permutasi, masalah dua sampel saling bebas, uji Wilcoxon-Mann-Whitney, mesin stripping.*

1. Pendahuluan

Salah satu tujuan digunakannya suatu analisis statistika adalah membuat kesimpulan tentang satu atau beberapa karakteristik tertentu dari satu atau beberapa populasi, baik dengan cara penaksiran ataupun pengujian hipotesis mengenai karakteristik tersebut. Salah satu analisis statistika tersebut adalah pengujian kesamaan dua rata-rata dari dua populasi yang saling bebas, yang sering disebut sebagai masalah dua sampel saling bebas. Dalam pengujian untuk masalah dua sampel saling bebas tersebut, masing-masing sampel harus diambil secara acak dari populasinya dan setiap pengamatan harus saling bebas satu sama lain.

Metode pengujian untuk masalah dua sampel saling bebas yang paling umum digunakan adalah uji t untuk kasus dua sampel saling bebas, dimana asumsi yang harus dipenuhi adalah data berasal dari populasi yang berdistribusi normal dan varians homogen. Akan tetapi dalam praktek data tidak selalu berasal dari populasi yang berdistribusi normal, misalnya data pendapatan, data besarnya klaim asuransi, data waktu operasi mesin, dan data waktu perawatan mesin.

Berdasarkan uraian di atas dibutuhkan suatu metode pengujian untuk masalah dua sampel saling bebas jika data yang dianalisisnya tidak berdistribusi normal. Hogg dan Craig (1995) mengemukakan suatu metode untuk masalah dua sampel saling bebas jika datanya tidak berdistribusi normal. Metode tersebut dikenal sebagai uji Wilcoxon-Mann-Whitney. Kelemahan uji Wilcoxon-Mann-Whitney adalah analisisnya tidak berdasarkan pada nilai datanya langsung tetapi pada peringkat dari datanya. Jika analisisnya didasarkan pada peringkat dari datanya bukan dari nilai datanya langsung, maka ada informasi yang dikandung oleh data yang tidak dilibatkan dalam analisis.

Berdasarkan uraian di atas, dapat diidentifikasi masalah yaitu dibutuhkannya suatu pengujian untuk masalah dua sampel saling bebas jika datanya tidak berdistribusi normal yang analisisnya didasarkan pada nilai datanya langsung. Adapun tujuan dari penulisan makalah ini adalah memperkenalkan metode pengujian untuk masalah dua sampel saling bebas jika datanya tidak berdistribusi normal yang analisisnya didasarkan pada nilai datanya langsung. Metodenya dikenal sebagai uji permutasi yang dikenalkan oleh R. A. Fisher tahun 1930 (Efron dan Tibshirani, 1993). Dalam analisisnya, metode ini menyusun kembali data dalam seluruh kombinasi yang mungkin dengan menggunakan teknik resampling permutasi.

Untuk ukuran sampel besar, pendekatan normal dari beberapa statistik uji permutasi sering terpenuhi (Johnson & Kotz (1985) dalam Hajarisman dan Saefuddin (2002)). Metode di atas akan diaplikasikan untuk data waktu perawatan mesin *stripping* merk Chuan Yung dan Ganson yang digunakan pada produksi obat-obatan di LAFI-DITKES AD (Lembaga Farmasi – Direktorat Kesehatan Angkatan Darat) Bandung Jawa Barat pada periode Maret 2004 sampai April 2005.

2. Masalah Dua Sampel Saling Bebas

Banyak penelitian yang memerlukan perbandingan antara dua keadaan atau tepatnya dua populasi. Misalnya membandingkan dua cara mengajar, dua cara produksi, daya sembuh dua macam obat dan lain sebagainya. Untuk keperluan ini akan digunakan dasar distribusi sampling mengenai selisih statistik, misalnya selisih rata-rata.

Misalkan ada dua populasi normal masing-masing dengan rata-rata μ_1 dan μ_2 sedangkan simpangan bakunya σ_1 dan σ_2. Secara independen dari populasi kesatu diambil sebuah sampel acak berukuran n_1 sedangkan dari populasi kedua sebuah sampel acak berukuran n_2. Dari kedua sampel ini berturut-turut didapatkan rata-rata dan simpangan bakunya, yaitu $\bar{x}_1, s_1$ dan $\bar{x}_2, s_2$. Akan diuji tentang rata-rata μ_1 dan μ_2.

Pasangan hipotesis nol dan tandingannya yang akan diuji adalah:

Uji dua pihak

$$\begin{cases} \mathrm{H}_0 : \mu_1 = \mu_2 \\ \mathrm{H}_1 : \mu_1 \neq \mu_2 \end{cases}$$

Uji pihak kanan

$$\begin{cases} \mathrm{H}_0 : \mu_1 = \mu_2 \\ \mathrm{H}_1 : \mu_1 > \mu_2 \end{cases}$$

Uji pihak kiri

$$\begin{cases} \mathrm{H}_0 : \mu_1 = \mu_2 \\ \mathrm{H}_1 : \mu_1 < \mu_2 \end{cases}$$

Dalam bagian ini hanya akan dibahas dua kasus. Kasus pertama ketika $\sigma_1 = \sigma_2 = \sigma$ dan σ diketahui. Kasus kedua ketika $\sigma_1 = \sigma_2 = \sigma$ dan σ tidak diketahui.

2.1. Kasus $\sigma_1 = \sigma_2 = \sigma$ dan σ diketahui

Statistik uji yang digunakan jika H_0 benar, adalah:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{1}$$

Dengan taraf arti α, kriteria pengujian adalah: (i) untuk uji dua pihak, tolak H_0 jika $|z| > z_{1-\alpha/2}$, dimana $z_{1-\alpha/2}$ adalah kuantil $(1 - \alpha/2)$ dari distribusi normal baku dengan; (ii) untuk uji pihak kanan, tolak H_0 jika $z > z_{1-\alpha}$, dimana $z_{1-\alpha}$ adalah kuantil $(1 - \alpha)$ dari distribusi normal baku; (iii) untuk uji pihak kiri, tolak H_0 jika $z < - z_{1-\alpha}$,.

2.2. Kasus $\sigma_1 = \sigma_2 = \sigma$ tetapi σ tidak diketahui

Jarang sekali σ_1 dan σ_2 diketahui besarnya. Jika H_0 benar dan $\sigma_1 = \sigma_2 = \sigma$ sedangkan σ tidak diketahui harganya, statistik yang digunakan adalah

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{2}$$

dimana

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \tag{3}$$

Menurut teori distribusi sampling, statistik t di atas berdistribusi t-Student dengan derajat kebebasan dk = (n_1 + n_2 - 2). Dengan taraf arti α, kriteria pengujian adalah: (i) untuk uji dua pihak, tolak H_0 jika $|t| > t_{1-\alpha/2}$, dimana $t_{1-\alpha/2}$ adalah kuantil $(1 - \alpha/2)$ dari distribusi t-Student dengan derajat kebebasan dk = (n_1 + n_2 - 2); (ii) untuk uji pihak kanan, tolak H_0 jika $t > t_{1-\alpha}$, dimana $t_{1-\alpha}$ adalah kuantil $(1 - \alpha)$ dari distribusi t-Student dengan derajat kebebasan dk = (n_1 + n_2 - 2); (iii) untuk uji pihak kiri, tolak H_0 jika $t < -t_{1-\alpha}$.

3. Uji Permutasi untuk Masalah Dua Sampel Saling Bebas

Uji permutasi merupakan sebuah teknik komputasi statistika yang mendahului lahirnya komputer. Gagasan ini telah diperkenalkan oleh R.A. Fisher pada tahun 1930. Kemampuan komputasi modern membuat uji permutasi praktis untuk digunakan.

Salah satu aplikasi dari uji permutasi adalah untuk masalah dua sampel saling bebas ketika data tidak berdistribusi normal. Misalkan diamati dua sampel acak saling bebas $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$ dan $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)$ yang mungkin diambil dari distribusi yang berbeda $\boldsymbol{F}$ dan $\boldsymbol{G}$.

Setelah mengamati $\boldsymbol{y}$ dan $\boldsymbol{z}$, akan diuji hipotesis H_0 tidak ada perbedaan antara $\boldsymbol{F}$ dan $\boldsymbol{G}$,

$$H0 : F = G \tag{4}$$

Statistik uji yang dipertimbangkan untuk pengujian di atas adalah,

$$\hat{\theta} = \bar{y} - \bar{z} \tag{5}$$

dimana

$$\bar{y} = \frac{\sum_{i=1}^{m} y_i}{m} \tag{6}$$

dan

$$\bar{z} = \frac{\sum_{i=1}^{n} z_i}{n} \tag{7}$$

Taraf signifikansi yang dicapai (*achieved significance level* - ASL) dari pengujian hipotesis pada Persamaan (4) didefinisikan sebagai berikut: (i) untuk uji pihak kanan, $\mathrm{ASL} = \mathrm{P}(\hat{\theta}^* \geq \hat{\theta})$; (ii) untuk uji pihak kiri; $\mathrm{ASL} = \mathrm{P}(\hat{\theta}^* \leq \hat{\theta})$; (iii) untuk uji dua pihak; $\mathrm{ASL} = \mathrm{P}(|\hat{\theta}^*| \geq |\hat{\theta}|)$. Peubah acak $\hat{\theta}^*$ mempunyai distribusi seperti distribusi pada hipotesis nol. Untuk taraf arti α tertentu, tolak hipotesis nol jika nilai ASL lebih kecil dari α.

Misalkan N adalah ukuran sampel gabungan, $N = m + n$, dan misalkan $\boldsymbol{v} = (v_1, v_2, \ldots, v_N)$ adalah vektor yang berisikan nilai-nilai data pengamatan gabungan. Misalkan juga $\boldsymbol{g} = (g_1, g_2, \ldots, g_N)$ merupakan vektor yang menunjukkan kelompok atau sampel untuk setiap data pengamatan. Vektor $\boldsymbol{g}$ mengandung n buah z dan m buah y. Ada sebanyak

$$\binom{N}{n} = \frac{N!}{n!m!} \tag{8}$$

vektor ***g*** yang mungkin, berkaitan dengan semua cara yang mungkin dalam mempartisi *N* elemen ke dalam dua himpunan bagian berukuran *m* dan *n*.

Misalkan g^* merupakan salah satu dari $\binom{N}{n}$ vektor yang mungkin dari *n* buah *z* dan *m* buah *y*, dan definisikan replikasi permutasi dari $\hat{\theta}$ sebagai, $\hat{\theta}^*$. Ada $\binom{N}{n}$ replikasi permutasi $\hat{\theta}^*$. Nilai ASL permutasinya adalah:

(i) untuk uji pihak kanan,

$$\text{ASL}_{\text{perm}} = \frac{\text{banyaknya}\{\hat{\theta}^* \geq \hat{\theta}\}}{\binom{N}{n}} \tag{9}$$

(ii) untuk uji pihak kiri,

$$\text{ASL}_{\text{perm}} = \frac{\text{banyaknya}\{\hat{\theta}^* \leq \hat{\theta}\}}{\binom{N}{n}} \tag{10}$$

(iii) untuk uji dua pihak,

$$\text{ASL}_{\text{perm}} = \frac{\text{banyaknya}\{|\hat{\theta}^*| \geq |\hat{\theta}|\}}{\binom{N}{n}} \tag{11}$$

Dalam praktek, ASL_{perm} biasanya didekati oleh metode Monte Carlo, sesuai dengan algoritma berikut ini.

1. Pilih B buah vektor saling bebas $g_1^*, g_2^*, \ldots, g_B^*$ masing-masing mengandung *n* buah *z* dan *m* buah *y* yang dipilih secara acak dari kumpulan $\binom{N}{n}$ vektor yang mungkin [*B* biasanya paling kecil 1000].
2. Evaluasi replikasi permutasi dari $\hat{\theta}$ yaitu

$$\hat{\theta}^* = \bar{y}^* - \bar{z}^* \tag{12}$$

3. Dekati nilai ASLperm oleh:

(i) untuk uji pihak kanan,

$$\text{ASL}_{\text{perm}} = \frac{\text{banyaknya}\{\hat{\theta}^* \geq \hat{\theta}\}}{B} \tag{13}$$

(ii) untuk uji pihak kiri,

$$\text{ASL}_{\text{perm}} = \frac{\text{banyaknya}\{\hat{\theta}^* \leq \hat{\theta}\}}{B} \tag{14}$$

(iii) untuk uji dua pihak,

$$\text{ASL}_{\text{perm}} = \frac{\text{banyaknya}\{|\hat{\theta}^*| \geq |\hat{\theta}|\}}{B} \tag{15}$$

4. Bahan

Bahan yang digunakan untuk mengaplikasikan metode yang dibahas dalam makalah ini adalah data sekunder mengenai waktu perawatan dua merk mesin *stripping* yang digunakan pada produksi obat-obatan di LAFI-DITKES AD (Lembaga Farmasi – Direktorat Kesehatan Angkatan Darat) Bandung Jawa Barat selama periode Maret 2004 sampai Juli 2005. Kedua merk mesin *stripping* tersebut adalah Chuan Yung dan Ganson. Waktu perawatan mesin yang dimaksud adalah selang waktu antara saat mesin mengalami kegagalan dan saat mesin selesai diperbaiki. Data di atas diperoleh dari Bagian Instalasi HAR di instansi tersebut. Tabel 1 menyajikan data waktu perawatan kedua merk tersebut (dalam jam).

Tabel 1. Data Waktu Perawatan (jam) Mesin Stripping Merk Chuan Yung dan Ganson

No	Chuan Yung	Ganson
1	96	144
2	96	72
3	144	120
4	168	240
5	48	
6	24	

Sumber: Instalasi HAR LAFI-DITKESAD tahun 2005

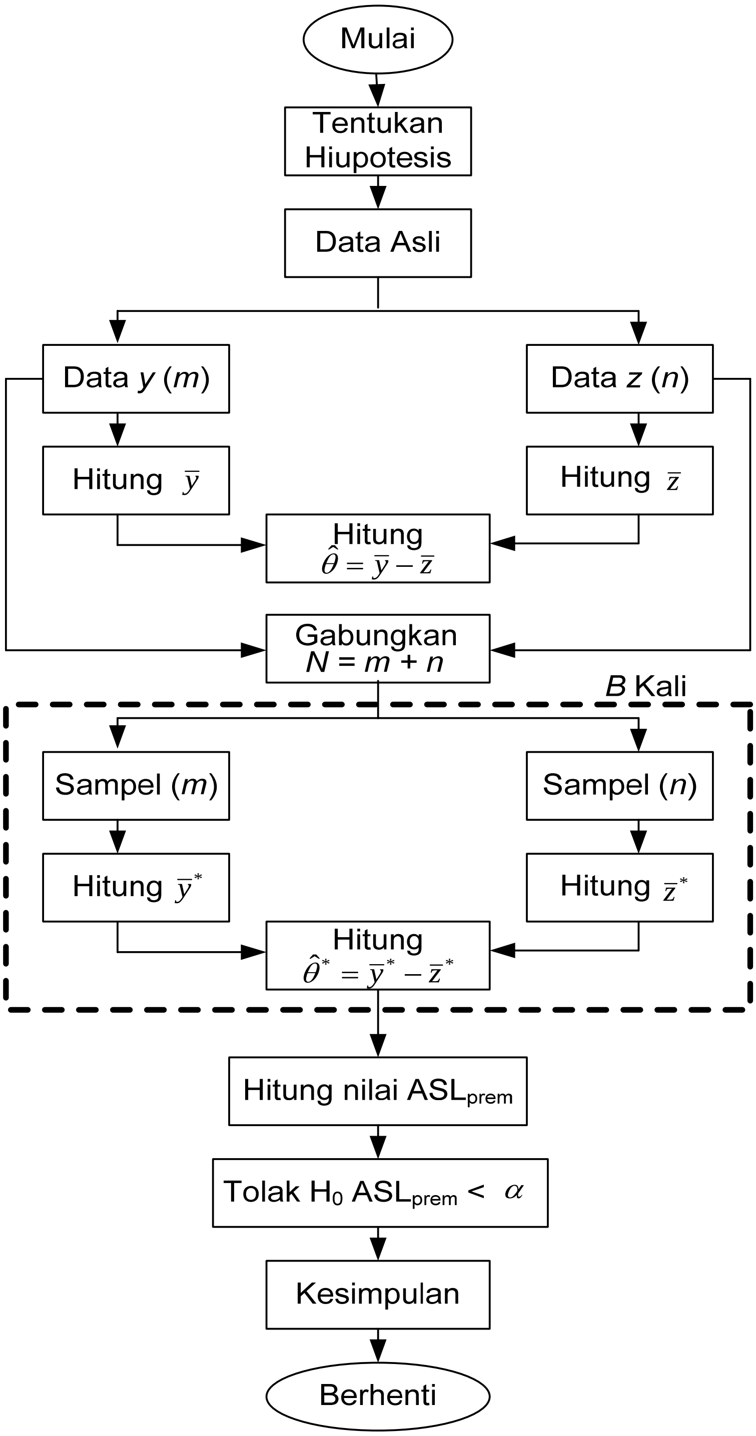
5. Metode

Pada bagian ini akan dibahas mengenai langkah-langkah uji permutasi untuk masalah dua sampel saling bebas. Langkah-langkah tersebut adalah sebagai berikut:

1. Nyatakan hipotesisnya,
   $\mathrm{H}_0 : \mu_1 = \mu_2$; tidak ada perbedaan rata-rata μ_1 & μ_2 ,
   $\mathrm{H}_1 : \mu_1 > \mu_2$ ($\mu_1 < \mu_2$ atau $\mu_1 \neq \mu_2$).
2. Hitung rata-rata untuk data pada sampel pertama (berukuran m), $\bar{y}$ dengan menggunakan Persamaan (6).
3. Hitung rata-rata untuk data pada sampel kedua (berukuran n), $\bar{z}$ dengan menggunakan Persamaan (7).
4. Hitung nilai perbedaan rata-rata untuk data pada sampel pertama dan kedua, $\hat{\theta}$, dengan menggunakan Persamaan (5).
5. Gabungkan data sampel pertama (berukuran *m*) dan kedua (berukuran *n*) untuk mendapatkan data sampel gabungan berukuran *N* = *m* + *n* pengamatan.
6. Ambil sebuah sampel secara acak berukuran *m* tanpa pengembalian dari data sampel gabungan untuk membentuk kelompok data pertama (y^*); *n* pengamatan sisanya merupakan kelompok data kedua (z^*).
7. Hitung rata-rata untuk kelompok data pertama $\left(\bar{y}^*\right)$ dan rata-rata untuk kelompok data kedua $\left(\bar{z}^*\right)$.
8. Hitung nilai perbedaan rata-rata untuk kelompok data pertama dan kedua, $\hat{\theta}^*$, dengan menggunakan Persamaan (12).
9. Ulangi Langkah 6 – 8 sebanyak B kali sehingga diperoleh nilai perbedaan sebanyak *B* buah yaitu $\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_B^*$.
10. Hitung nilai ASL permutasi:
    (1) Untuk uji pihak kanan, gunakan Persamaan (13).
    (2) Untuk uji pihak kiri, gunakan Persamaan (14).
    (3) Untuk uji dua pihak, gunakan Persamaan (15).

11. Untuk taraf arti α tertentu, tolak hipotesis nol jika nilai ASL permutasi lebih kecil dari α.

Gambar 1 di bawah ini mengilustrasikan diagram alir untuk metode di atas.



Gambar 1. Diagram Alir Metode Perhitungan Nilai ASLprem Uji Permutasi untuk Masalah Dua Sampel Saling Bebas

6. Hasil Perhitungan

Berikut ini akan dilakukan uji permutasi untuk dua merk mesin *stripping* yang mengalami waktu perawatan di LAFI-DITKES AD (Lembaga Farmasi – Direktorat Kesehatan Angkatan Darat) Bandung Jawa Barat. Pernyataan hipotesisnya adalah:

$H_0 : \mu_1 = \mu_2$; rata-rata waktu perawatan mesin *stripping* merk Chuan Yung sama dengan rata-rata waktu perawatan mesin *stripping* merk Ganson.

$H_1 : \mu_1 < \mu_2$ rata-rata waktu perawatan mesin *stripping* merk Chuan Yung lebih kecil dibanding rata-rata waktu perawatan mesin *stripping* merk Ganson.

Nilai statistik uji permutasi untuk kedua data waktu perawatan mesin *stripping* dihitung menggunakan Persamaan (5), yaitu:

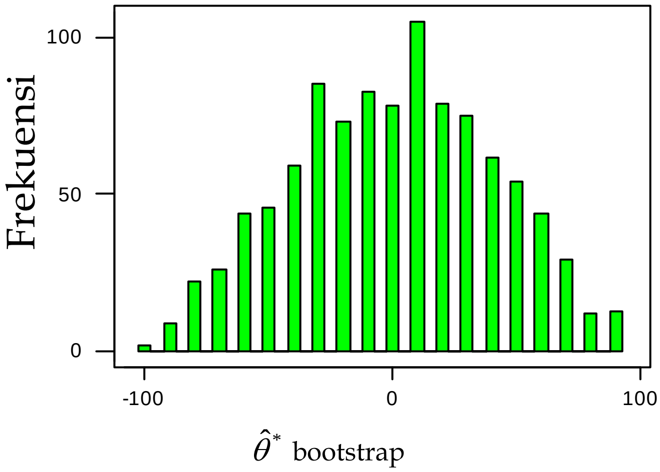
$$\begin{aligned}\hat{\theta} &= \bar{y} - \bar{z} \\ &= 96 - 144 \\ &= -48\end{aligned}$$

Dengan bantuan program makro Minitab yang dioperasikan pada perangkat lunak Minitab 13.20 (programnya disajikan pada Lampiran 3). Untuk mempermudah penghitungan nilai statistik uji permutasi $\hat{\theta}^*$ dan penggabungan data pertama dan kedua sebanyak 10 pengamatan, kemudaian ambil sampel secara acak tampa pengembalian dari data pertama sebanyak berukuran (6) termasuk pada kelompok pertama dan sisanya (4) termasuk kepada sampel kelompok kedua dilakukan berulang-ulang sebanyak 1000 kali untuk mendapatkan nilai statistik uji permutasi $\hat{\theta}^*$ bootstrap. Tabel 2 menyajikan nilai statistik uji permutasi $\hat{\theta}^*$ tersebut.

Tabel 2. Nilai Statistik Uji Permutasi $\hat{\theta}^*$ untuk Setiap Sampel Bootstrap

Nomor Sampel Bootstrap	Statistik Uji Permutasi ($\hat{\theta}^*$)
1	22
2	32
3	-38
⋮	⋮
1000	-8

Sedangkan nilai statistik uji permutasi $\hat{\theta}^*$ dari hasil pengulangan untuk data sampel bootstrap berjumlah 1000. Bentuk histogram dari 1000 nilai statistik uji permutasi $\hat{\theta}^*$ bootstrap tersebut diilustrasikan dalam Gambar 2.



Gambar 2. Histogram Nilai Statistik Uji Permutasi $\hat{\theta}^*$ untuk Data Sampel Bootstrap

Dari hasil perhitungan, banyaknya nilai statistik uji permutasi $\hat{\theta}^*$ untuk data sampel bootstrap yang lebih kecil dari nilai statistik uji permutasi $\hat{\theta}$ untuk data sampel asli ada sebanyak 103. jadi nilai ASL permutasi dapat dihitung dengan menggunakan Persamaan (14), yaitu:

$$\mathrm{ASL}_{\mathrm{perm}} = \frac{103}{1000} = 0{,}103 \approx 10\%$$

Nilai tersebut menunjukan bahwa kemungkinan salah apabila hipotesis H_0 ditolak (rata-rata waktu perawatan mesin *stripping* merk Chuan Yung berbeda dengan rata-rata waktu perawatan mesin *stripping* merk Ganson) adalah sebesar 10%.

Jika taraf signifikansi yang digunakan adalah α = 0,05 maka H_0 diterima. Karena nilai ASL permutasinya lebih kecil dari nilai taraf signifikansi tersebut, jadi disimpulkan bahwa rata-rata waktu perawatan mesin *stripping* merk Chuan Yung sama dengan rata-rata waktu perawatan mesin *stripping* merk Ganson di LAFI-DITKES AD (Lembaga Farmasi – Direktorat Kesehatan Angkatan Darat) Bandung Jawa Barat.

7. Kesimpulan

Uji permutasi yang dikenalkan oleh R. A. Fisher tahun 1930 merupakan salah satu alternatif untuk pengujian masalah dua sampel saling bebas jika datanya bukan berasal dari populasi yang berdistribusi normal. Dalam analisisnya, metode ini didasarkan pada datanya langsung. Untuk ukuran sampel besar, pendekatan normal dari beberapa statistik uji permutasi sering terpenuhi.

Berdasarkan hasil perhitungan dengan menerapkan teori yang dibahas terhadap data waktu perawatan dua merk mesin *stripping* yang digunakan pada produksi obat-obatan di LAFI-DITKES AD (Lembaga Farmasi – Direktorat Kesehatan Angkatan Darat) Bandung Jawa Barat selama periode Maret 2004 sampai Juli 2005, dapat disimpulkan bahwa Nilai ASL permutasi untuk data di atas sebesar 0,103 ≈ 10%. Nilai tersebut menunjukan bahwa kemungkinan salah apabila hipotesis H_0 ditolak (rata-rata waktu perawatan mesin *stripping* merk Chuan Yung berbeda dengan rata-rata waktu perawatan mesin *stripping* merk Ganson) adalah sebesar 10%. Dengan taraf signifikansi α = 5% disimpulkan bahwa rata-rata waktu perawatan mesin *stripping* merk Chuan Yung sama dengan rata-rata waktu perawatan mesin *stripping* merk Ganson di LAFI-DITKES AD (Lembaga Farmasi – Direktorat Kesehatan Angkatan Darat) Bandung Jawa Barat.

Daftar Pustaka

[1]. Efron, B., dan Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York.

[2]. Hajarisman, N. dan Saefuddin, A. (2002). Pendekatan Nonparametrik Untuk Analisis Trand pada Respon Biner. *Disampaikan pada Seminar Nasional Statistika, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, pada tanggal 28 September 2002*.

[3]. Hogg, R. V. and Craig, A. T. (1995). *Introduction to Mathematical Statistics*. Fifth Edition. Macmillan Publishing Company, New York.

[4]. Sudjana. (1996). *Metoda Statistika*. Penerbit Tarsito. Bandung.

[5]. Walpole, R. E., dan Myers, R. H. (1995). *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*. Penerbit ITB. Bandung.