## BAB III

## UJI PERMUTASI ASL UNTUK PENGUJIAN HIPOTESIS

Pengujian hipotesis merupakan salah satu aspek statistik induktif yang sangat berguna, karena banyak jenis persoalan keputusan dapat diformulasikan sebagai masalah pengujian hipotesis. Salah satu persoalan keputusan yang selanjutnya akan dibahas pada tugas akhir ini adalah permasalahan dua sampel random independen.

Dalam mengambil keputusan, sebelumnya harus dibuat suatu pemisalan (asumsi) tentang populasinya. Pemisalan tersebut yang bisa bernilai benar atau tidak benar disebut sebagai hipotesis statistik. Keputusan menerima atau menolak hipotesis nol didasarkan pada pengujian statistik yang dihitung dari data dalam sebuah sampel random.

### 3.1 Hipotesis Tingkat Signifikan yang Tercapai

Jika data sampel random tersebut sangat berbeda dibandingkan dengan jika hipotesis itu benar, maka dikatakan beda itu *nyata ( significant )* , sehingga cenderung dilakukan penolakan terhadap hipotesis nol.

Dua jenis error bisa dibuat pada saat pengujian hipotesis ( Tabel 3.1 ).

Jika hipotesis nol ditolak ketika hipotesis nol benar, maka error tipe I telah dilakukan. Jika hipotesis nol diterima ketika hipotesis nol salah, maka error tipe II telah dilakukan.

Probabilita – probabilita terjadinya error tipe I dan tipe II diberi simbol khusus sebagai berikut :

$$\alpha = P\ \{\text{ error tipe I }\} = P\ \{\text{ menolak } H_0 \mid H_0 \text{ benar }\}$$

$\beta = P$ { error tipe II } = P { menerima $H_0$ | $H_0$ salah }

Karena ini adalah kesalahan, maka dalam pengujian hipotesis harus diusahakan untuk membuat nilai kesalahan itu kecil.

Tabel 3.1 Keputusan dalam pengujian hipotesis.

| **Kriteria Uji Sampel Jatuh ke dalam** | **Keputusan** | **Data berasal dari populasi yang** | |
|---|---|---|---|
| | | **$H_0$ benar, $H_1$ salah** | **$H_0$ salah, $H_1$ benar** |
| Wilayah penerimaan Uji Tidak Nyata | Menerima $H_0$ Menolak $H_1$ | Keputusan yang benar Peluang harus tinggi Lambang: $1-\alpha$ = koefisien kepercayaan | Keputusan yang salah Error Jenis II Peluang harus kecil Lambang : $\beta$ |
| Wilayah penolakan Uji Nyata | Menolak $H_0$ Menerima $H_1$ | Keputusan yang salah Error Jenis I Peluang harus kecil Lambang : $\alpha$ = taraf nyata | Keputusan yang benar Peluang harus tinggi Lambang: $1-\beta$ = kuasa uji |

Jadi , $\alpha$ = P (kesalahan jenis I ) = P ( menolak $H_0$ | $H_0$ benar ) harus kecil, dan juga $\beta$ = P (kesalahan jenis II ) = P ( menerima $H_0$ | $H_0$ salah ) harus kecil.

Probabilita error tipe I, $\alpha$ , disebut juga **taraf nyata ( significant level )**. Biasanya probabilita error tipe I dikontrol oleh lokasi daerah kritis. Probabilita kesalahan penolakan $H_0$ secara langsung dikontrol oleh pembuat keputusan. Dan penolakan $H_0$ selalu sebuah kesimpulan yang kuat.

Pada umumnya $\alpha$ ditentukan dahulu, misalnya 0,05 ; 0,01 ; atau 0,001 ; dan antara semua pengujian dengan $\alpha$ akan dipilih satu pengujian yang meminimumkan $\beta$. Misal $\alpha = 0{,}05$ ; berarti dalam 100 kali, dilakukan penolakan $H_0$ sebanyak 5 kali padahal $H_0$ benar; atau mempunyai kepercayaan 95 % bahwa $H_0$ benar. Dapat juga dikatakan $H_0$ ditolak dengan **taraf nyata ( significant level )** sebesar 5 %.

Dengan mengamati statistik uji , pengujian tingkat signifikan yang tercapai, akan memberi informasi mengenai probabilitas pengamatan ketika hipotesis nol bernilai benar. Tingkat signifikan yang tercapai dalam bahasa Inggris dinyatakan dengan **achieved significance level** atau disingkat **ASL**. Misalkan akan dilakukan pengamatan pada statistik uji $\hat{\theta}$ , maka dapat dinyatakan **ASL** sebagai berikut :

$$\mathbf{ASL} = \mathrm{Prob}_{H_1} \{ \hat{\theta}^* \geq \hat{\theta} \} \qquad \text{.........(3.1)}$$

Suatu pengujian hipotesis dari $H_0$ terdiri dari penghitungan **ASL** dan memperhatikan nilainya apakah cukup kecil dengan suatu ambang konvensi tertentu.

Secara formal dapat dipilih sebuah probabilitas kecil $\alpha$ , seperti 0,05 atau 0,01 dan menolak $H_0$ jika **ASL** kurang dari $\alpha$. Jika **ASL** lebih besar dari $\alpha$, maka diterima $H_0$.

Untuk mengamati **ASL** dan mengukur kuatnya bukti menolak /melawan $H_0$ , dapat ditentukan dengan suatu konvensi kasar. Penentuan besarnya $\alpha$ tergantung oleh peneliti yang akan mengambil suatu keputusan.

Misalkan digunakan konvensi seperti berikut :

| **ASL** $< \alpha$ | KETERANGAN |
|---|---|
| **ASL** $< 0{,}10$ | $H_0$ dapat ditolak |

(3.2)

Setelah menentukan ASL yang akan digunakan, maka langkah selanjutnya adalah menghitung besar ASL pada suatu data.

Sebuah uji hipotesis tradisional untuk suatu data, dapat dimulai dengan asumsi bahwa F dan G merupakan distribusi normal dengan kemungkinan rata-rata yang berbeda.

$$F \sim N(\mu_T, \sigma^2) \quad G \sim N(\mu_C, \sigma^2)$$

Karena hipotesis nol menyatakan distribusi F dan G yang tidak berbeda, maka hipotesis nol-nya dapat dinyatakan sebagai berikut :

$$H_0 : \mu_T = \mu_C$$

Di bawah $H_0$ :

$$\begin{aligned} E(\bar{z} - \bar{y}) &= E\,\bar{z} - E\,\bar{y} \\ &= \mu_T - \mu_C \\ &= 0 \end{aligned}$$

Sedangkan variannya

$$\begin{aligned} Var(\bar{z} - \bar{y}) &= Var\{\bar{z} + (-\bar{y})\} \\ &= Var\,\bar{z} + Var(-\bar{y}) \\ &= \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m} \end{aligned}$$

Diasumsikan bahwa $\sigma_1^2 = \sigma_2^2 = \sigma^2$ , sehingga dapat dinyatakan :

$$Var(\bar{z} - \bar{y}) = \sigma^2\left(\tfrac{1}{n} + \tfrac{1}{m}\right).$$

Sehingga di bawah $H_0$ , $\hat{\theta} = \bar{z} - \bar{y}$ mempunyai sebuah distribusi normal dengan rata-rata 0 dan varian $\sigma^2\left(\tfrac{1}{n} + \tfrac{1}{m}\right)$ ,

$$H_0 : \hat{\theta} \sim N\left(0, \sigma^2\left(\tfrac{1}{n} + \tfrac{1}{m}\right)\right); \quad \ldots\ldots\ldots\ldots\ldots(3.3)$$

Menurut dalil limit sentral, jika $x_1$ dan $x_2$ , dalam kasus ini z dan y, adalah normal maka distribusi $\bar{z} - \bar{y}$ juga normal. Sehingga statistiknya adalah :

$$Z = \frac{\hat{\theta}}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}}$$

$$= \frac{\hat{\theta}}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

Dengan mengingat bahwa $\hat{\theta}^*$ mempunyai distribusi hipotesis nol, yaitu distribusi $\hat{\theta}$ jika $H_0$ benar, dan dari melakukan pengamatan pada $\hat{\theta}$ secara aktual, maka **ASL** adalah probabilitas dari suatu variabel random $\hat{\theta}^*$ yang didistribusikan seperti pada (3.3) yang melampaui $\hat{\theta}$ sebagai berikut :

$$\textbf{ASL} = \mathrm{Prob}_{H_0}\{\hat{\theta}^* \geq \hat{\theta}\}$$

$$\textbf{ASL} = \mathrm{Prob}\left\{ Z > \frac{\hat{\theta}}{\sigma\sqrt{\tfrac{1}{n} + \tfrac{1}{m}}} \right\}$$

$$\textbf{ASL} = 1 - \phi\left( \frac{\hat{\theta}}{\sigma\sqrt{\tfrac{1}{n} + \tfrac{1}{m}}} \right) \quad \ldots\ldots\ldots\ldots\ldots(3.4)$$

dimana $\phi$ adalah fungsi distribusi kumulatif dari variasi standar normal Z. Karena $\sigma$ tidak diketahui , maka standar estimasinya merupakan gabungan dari $\sigma$ yang diambil dari dua grup yang berbeda yaitu

$$\bar{\sigma}_1 = \left\{ \left[ \sum_{i=1}^{n} (Z_i - \bar{Z})^2 \right] \Big/ [n-1] \right\}^{1/2}$$

$$\bar{\sigma}_2 = \left\{ \left[ \sum_{i=1}^{m} (Y_i - \bar{Y})^2 \right] \Big/ [m-1] \right\}^{1/2}$$

adalah :

$$\bar{\sigma} = \left\{ \left[ \sum_{i=1}^{n} (Z_i - \bar{Z})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2 \right] \Big/ [n+m-2] \right\}^{1/2}$$

Perhitungan ini memperlakukan $\bar{\sigma}$ seperti layaknya suatu konstanta. Dengan menggunakan uji Student-t yang memperhatikan akan kerandoman dalam $\bar{\sigma}$ , dan dengan statistik uji $\dfrac{\hat{\theta}}{\left\{ \left[ \sum_{i=1}^{n} (Z_i - \bar{Z})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2 \right] \Big/ [n+m-2] \right\}^{1/2} \sqrt{1/n + 1/m}}$ , akan memberikan hasil sebagai berikut :

$$ASL = \text{Prob}\left\{ t_{n+m-2} > \frac{\hat{\theta}}{\left\{ \left[ \sum_{i=1}^{n} (Z_i - \bar{Z})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2 \right] \Big/ [n+m-2] \right\}^{1/2} \sqrt{1/n + 1/m}} \right\} \quad \ldots(3.5)$$

$t_{n+m-2}$ mengindikasikan suatu variasi t dengan $n + m - 2$ derajat kebebasan. Statistik ini mempunyai distribusi $t_{n+m-2}$ di bawah hipotesis nol.

## 3.2 Uji Permutasi ASL

Berikut ini akan disajikan konsep mengenai uji permutasi ASL dalam pengujian hipotesis yang kemudian akan diterapkan pada suatu studi kasus.

### 3.2.1 Uji Permutasi

Uji permutasi didasarkan pada penyajian statistik order dari pengamatan suatu data $x = (z, y)$ dari suatu permasalahan dua sampel.

Aplikasi penting dari uji permutasi adalah pada permasalahan dua sampel random independen yang diambil dari distribusi probabilitas yang berbeda F dan G.

$F \rightarrow \quad Z = (Z_1, Z_2, \ldots, Z_n)$ .......(3.6)

$G \rightarrow \quad Y = (Y_1, Y_2, \ldots, Y_n)$ ........(3.7)

dimana F dan G adalah bebas.

Dalam pengamatan tersebut dilakukan pengujian hipotesis nol ($H_0$), dimana $H_0$ menyatakan bahwa F dan G mempunyai probabilitas yang sama untuk semua himpunan. Uji hipotesis $H_0$ tersebut dapat ditulis sebagai berikut :

$H_0 : F = G$ ............(3.8)

$Prob_F \{A\} = Prob_G \{A\}$ ............(3.9)

dengan A merupakan sebarang subset dari ruang sampel yang umum dari z dan y. Jika $H_0$ benar, maka tidak ada perbedaan antara sifat probabilitas dari sebuah random z atau random y.

Selain itu diformulasikan pula hipotesis alternatifnya $H_1$ , yaitu dalam hal bukti-bukti yang dikumpulkan tidak menunjang $H_0$. Sehingga dari hipotesis nol di atas dapat dinyatakan :

$H_1 : F \neq G$ ............(3.10)

$Prob_F \{A\} \neq Prob_G \{A\}$ ............(3.11)

Kemudian dapat ditentukan nilai ASL dengan mendasarkan pada seriusnya kesalahan akibat menolak $H_0$ padahal $H_0$ benar; dengan demikian peluang menerima $H_0$ bila $H_0$ memang benar adalah $1 - \alpha$. Secara sekaligus, mungkin saja $H_0$ diterima padahal $H_1$ yang benar. Sehingga bila disebut dengan $\beta$ , maka peluang menerima $H_1$ bila $H_1$ memang benar adalah $1 - \beta$.

Dalam melakukan sederetan algoritma pada pengerjaan uji Permutasi **ASL** diperlukan suatu teknik-teknik yang mendukung penyelesaiannya. Sehingga sebelum diolah lebih lanjut seluruh data telah disusun dan diurutkan dari nilai terkecil ke nilai terbesar. Dan tiap data mempunyai keterangan grupnya. Untuk keterangan mengenai grup data, dinyatakan dengan "z" untuk Perlakuan atau "y" untuk Kontrol.

Misalnya penggunaan **Quick Sort** untuk pengurutan data dan **Merging** pada penggabungan datanya, sebelum dilakukan proses permutasi dengan himpunan ganda.

**Quick Sort** merupakan suatu algoritma pengurutan data yang mempergunakan teknik pemecahan data menjadi partisi-partisi, sehingga metode ini disebut juga dengan nama *partition exchange sort*. Secara garis besar metode ini dijelaskan sebagai berikut :

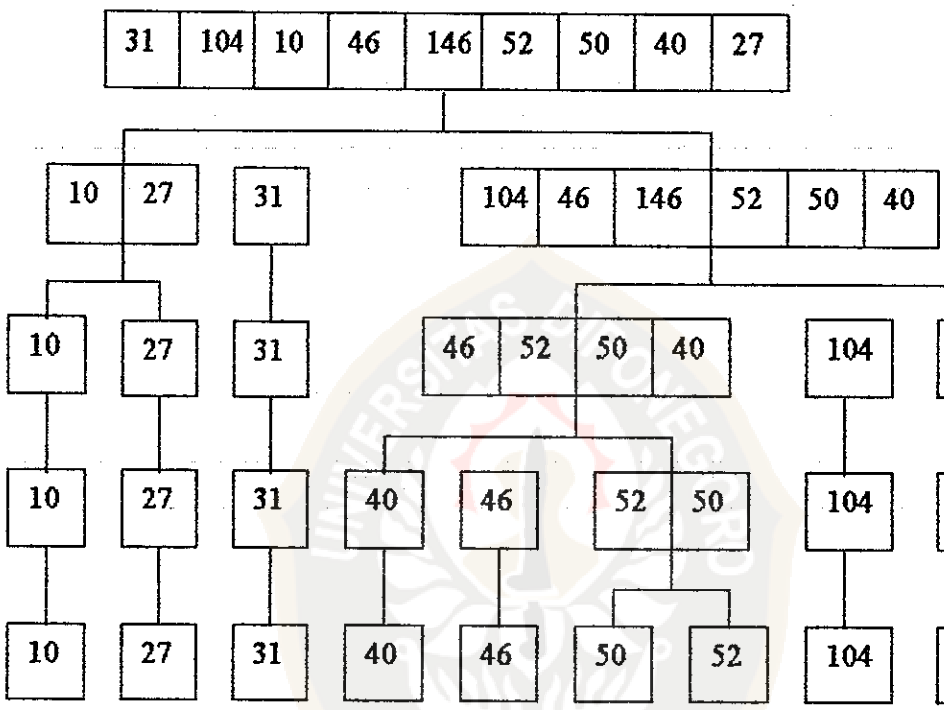
Misalkan akan diurutkan vektor A yang mempunyai N elemen.

- Untuk memulai iterasi pengurutan, pertama-tama sebuah elemen dipilih dari vektor tersebut, biasanya elemen pertama, sebut X.
- Menyusun semua elemen dengan menempatkan X pada posisi J sedemikian sehingga elemen ke 1 sampai ke J-1 mempunyai nilai lebih kecil dari X dan elemen ke J+1 sampai ke N mempunyai nilai lebih besar dari X. Dengan demikian terdapat dua buah subvektor, subvektor pertama nilai elemen lebih kecil dari X, subvektor kedua nilai elemennya lebih besar dari X.
- Proses diulang pada kedua subvektor, sehingga diperoleh empat subvektor.
- Proses di atas diulang pada setiap subvektor sehingga seluruh vektor semua elemennya menjadi terurutkan.

Berikut contoh metode ini yang digunakan pada vektor Control:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 31 | 104 | 10 | 46 | 146 | 52 | 50 | 40 | 27 |

Gambar 3.3 Ilustrasi metode **Quicksort**




**Merging** adalah proses penggabungan dua kumpulan data yang keduanya dalam keadaan urut menjadi satu kumpulan yang juga dalam keadaan urut. Secara umum proses **Merging** dijelaskan sebagai berikut : ( untuk mempermudah pemahaman digunakan dua buah vektor ). Dianggap bahwa dua buah larik yang akan digabung sudah dalam keadaan urut naik. Kedua buah vektor tersebut adalah sebagai berikut :

Vektor I ( Treatment ) :

| | | | | | | |
|---|---|---|---|---|---|---|
| 16 | 23 | 38 | 94 | 99 | 141 | 197 |

Vektor II ( Control ) :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 10 | 27 | 31 | 40 | 46 | 50 | 52 | 104 | 146 |

- Mengambil elemen pertama dari Vektor 1, ($V_1$), dan Vektor 2, ($V_2$).
- Membandingkan nilai kedua elemen ini. Jika $V_1 > V_2$ , maka $V_2$ dikopikan ke vektor ketiga; Jika tidak, $V_2$ dikopikan ke vektor ketiga.
- Untuk vektor yang elemennya dikopikan ke vektor ketiga, elemen berikutnya yang dibandingkan adalah elemen yang terletak pada subskrib berikutnya.

Dalam contoh di atas ( $V_1 = 16$ dan $V_2 = 10$ ) yang terjadi adalah $V_1 > V_2$ , sehingga $V_2$ dikopikan ke vektor ketiga. Untuk pembandingan berikutnya, digunakan $V_1 = 16$, $V_2 = 27$. Proses ini diulang terus sampai salah satu vektor habis dikopikan terlebih dahulu. Kemudian vektor satunya juga dikopikan ke vektor ketiga. Hasil **Merging** kedua vektor di atas adalah sebagai berikut ( yang digaris bawah berasal dari Vektor 2 ) :

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\underline{10}$ | 16 | 23 | $\underline{27}$ | $\underline{31}$ | 38 | $\underline{40}$ | $\underline{46}$ | $\underline{50}$ | $\underline{52}$ | 94 | 99 | $\underline{104}$ | 141 | $\underline{146}$ | 197 |

Kemudian diambil N sebagai sampel yang telah disusun dengan ukuran n + m, dan diambil $v = ( v_1, v_2, \ldots, v_N )$ sebagai vektor dari nilai yang telah disusun dan diurutkan. Juga diambil $g = ( g_1, g_2, \ldots, g_N )$ sebagai vektor yang mengindikasikan grup masing-masing observasi. Selanjutnya secara bersamaan v dan g menyajikan informasi yang sama seperti $x = ( z, y )$.

Vektor g terdiri dari z sejumlah n dan y sejumlah m. Sehingga ada

$$\binom{N}{n} = \frac{N\,!}{n!\, m!} \qquad \ldots\ldots \quad (3.13)$$

vektor g yang mungkin, berhubungan dengan semua cara yang mungkin dari pemartisian N elemen menjadi dua subset dari ukuran n dan m.

Menyusun suatu permutasi himpunan ganda untuk mendapatkan vektor-vektor v seperti di atas dapat dilakukan hampir sama dengan jika melakukan kombinasi. Sehingga bila akan dibentuk permutasi pada himpunan ganda dapat dilakukan dengan menyusun kombinasi pada salah satu grupnya dan grup lain merupakan sisa dari gabungan dengan kombinasi pertama yang telah tersusun tadi. Untuk melakukan kombinasi r dari gabungan $S = \{ S_1 , S_2 , \ldots, S_N \}$ dengan urutan leksikografik, maka digunakan langkah berikut ini :

- Dimulai dengan kombinasi $S_1 S_2 \ldots S_r$ .
- Jika kombinasi r $s_1 , \ldots, s_r$ telah dibentuk, jika $s_1 = n - r + 1$ , $s_2 = n - r + 2$ , … , $s_r = n$, maka semua kombinasi r telah selesai dibuat. Jika tidak maka pilih bilangan bulat j yang terbesar sehingga $s_j + 1 \leq n$ dan $s_j + 1$ bukan salah satu dari $s_1 , \ldots, s_r$. Kombinasi r berikutnya adalah $s_1 , \ldots, s_{j-1} , s_{j+1} , s_{j+2} , \ldots, s_{j+(r-j+1)}$.

Misal :

$X_1$ = | 10 16 23 27 31 38 40 46 50 52 94 99 104 141 146 197 |

$X_2$ = | 10 16 23 27 31 38 40 46 52 50 94 99 104 141 146 197 |

$X_3$ = | 10 16 23 27 31 38 40 46 94 52 50 99 104 141 146 197 |

∘
∘
∘

$X_i$ = | 10 16 23 27 31 38 40 46 197 52 94 99 104 141 146 50 |

$X_{i+1}$ = | 10 16 23 27 31 38 40 52 50 46 94 99 104 141 146 197 |

∘
∘
∘

$X_l$ = | 10 16 23 27 31 38 40 197 50 52 94 99 104 141 146 46 |

o

o

o

$$X_{\binom{N}{n}} = \boxed{46\ \ 50\ \ 52\ \ 94\ \ 99\ \ 104\ \ 141\ \ 146\ \ 197\ \ 10\ \ 16\ \ 23\ \ 27\ \ 31\ \ 38\ \ 40}$$

### 3.2.2 Statistik Penguji

Sebuah uji hipotesis dimulai dengan suatu statistik uji $\hat{\theta}$ sebagai selisih rata-rata, $\hat{\theta} = \bar{z} - \bar{y}$. Sehingga jika sampel pada Treatment mempunyai hasil lebih baik daripada Control pada eksperimen tersebut, maka diharapkan nilai $\hat{\theta} = \bar{z} - \bar{y}$ menjadi besar. Dalam menjalankan pengujian hipotesis ini tidak perlu mengkuantifikasikan arti kata "besar". Yang penting adalah bahwa semakin besar nilai $\hat{\theta}$ didapatkan, maka semakin kuat bukti melawan $H_0$.

Dengan mengamati $\hat{\theta}$, pengujian tingkat signifikan yang tercapai (*achieved significance level*) atau disingkat **ASL**, akan memberi informasi mengenai probabilitas pengamatan ketika hipotesis nol bernilai benar.

$$\textbf{ASL} = \mathrm{Prob}_{H_0}\{\hat{\theta}^* \geq \hat{\theta}\}$$

Kuantitas $\hat{\theta}$ pada pernyataan di atas ditentukan pada nilai pengamatannya, variabel random $\hat{\theta}^*$ mempunyai distribusi hipotesis nol, yaitu distribusi dari $\hat{\theta}$ jika $H_0$ benar. Notasi bintang membedakan antara pengamatan aktual $\hat{\theta}$ dengan suatu hipotetik $\hat{\theta}^*$ yang dibangun berdasarkan $H_0$.

Dengan asumsi bahwa F dan G merupakan distribusi normal dengan kemungkinan rata-rata yang berbeda, sebagai berikut :

$F \sim N(\mu_T, \sigma_1^2)$ ; $G \sim N(\mu_C, \sigma_2^2)$ dan $H_0 : \mu_T = \mu_C$ ,

dan dengan asumsi $\sigma_1^2 = \sigma_2^2 = \sigma^2$ , dapat dinyatakan bahwa di bawah $H_0$ , $\hat{\theta} = \bar{z} - \bar{y}$ mempunyai sebuah distribusi normal dengan rata-rata 0 dan varian $\sigma^2\left(\frac{1}{n} + \frac{1}{m}\right)$ ,

$$H_0 : \hat{\theta} \sim N\left(0, \sigma^2\left(\frac{1}{n} + \frac{1}{m}\right)\right).$$

Sehingga statistiknya adalah :

$$Z = \frac{\hat{\theta}}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

Karena $\sigma$ tidak diketahui , maka standar estimasinya adalah :

$$\bar{\sigma} = \left\{\left[\sum_{i=1}^{n}(Z_i - \bar{Z})^2 + \sum_{j=1}^{m}(Y_j - \bar{Y})^2\right] \Big/ [n + m - 2]\right\}^{1/2}$$

sehingga statistik uji yang digunakan adalah

$$\frac{\hat{\theta}}{\left\{\left[\sum_{i=1}^{n}(z_i - \bar{z})^2 + \sum_{j=1}^{m}(Y_j - \bar{Y})^2\right] \Big/ [n + m - 2]\right\}^{1/2} \sqrt{\frac{1}{n} + \frac{1}{m}}}$$

Dalam uji permutasi ASL, penentuan statistik uji dimulai dengan memperhatikan :

N = jumlah sampel yang telah disusun dengan ukuran n + m,

v = vektor dari nilai yang telah disusun dan diurutkan

$v = (v_1, v_2, \ldots, v_N)$

g = vektor yang mengindikasikan grup masing-masing observasi

$g = (g_1, g_2, \ldots, g_N)$ .

Vektor g terdiri dari z sejumlah n dan y sejumlah m. Sehingga ada

$$\binom{N}{n} = \frac{N!}{n!\,m!}$$

**Lema :**

Di bawah $H_0 : F = G$, vektor g mempunyai probabilitas $1\Big/\binom{N}{n}$ dari sebarang nilai-nilai yang mungkin.

**Bukti :**

Jika $H_0 : F = G$, maka probabilitas dari seluruh $\binom{N}{n} = \frac{N!}{n!\,m!}$ replikasi adalah $\sum \mathrm{Prob}_{H_0}\{\hat{\theta}_i^* \geq \hat{\theta}\} = 1$, sehingga tiap vektor g mempunyai probabilitas $\mathrm{Prob}_{H_0}\{\hat{\theta}^* \geq \hat{\theta}\}$, dimana tanda $^*$ memperlihatkan tiap replikasi permutasi yang terjadi , Sehingga $\mathrm{Prob}_{H_0}\{\hat{\theta}_i^* \geq \hat{\theta}\} = 1\Big/\binom{N}{n}$, dengan i = 1, 2, ...., $\binom{N}{n}$, yang merupakan replikasi permutasi ke i yang terjadi.

Hal ini bisa dijelaskan sebagai berikut bahwa seluruh vektor hasil permutasi dari z dan y adalah mempunyai kemungkinan yang sama jika F = G. Sebaliknya jika F ≠ G, maka suatu permutasi dari z dan y belum tentu mempunyai kemungkinan yang sama dengan permutasi lainnya.

Suatu statistik uji $\hat{\theta}$ sebagai fungsi dari g dan v dapat dinyatakan

$$\hat{\theta} = S(g, v) \qquad \ldots\ldots \quad (3.14)$$

Sehingga $\hat{\theta} = \bar{z} - \bar{y}$ dapat dinyatakan sebagai

$$\hat{\theta} = \frac{1}{n}\sum_{g_i = z} v_i - \frac{1}{m}\sum_{g_i = y} v_i \qquad \ldots\ldots \quad (3.15)$$

dengan $\sum_{g_i = z} v_i$ mengindikasikan suatu jumlah dari $v_I$ untuk $i = 1, 2, \ldots, N$ yang mempunyai $g_i = z$. Sedangkan $\sum_{g_i = y} v_i$ mengindikasikan suatu jumlah dari $v_I$ untuk $i = 1, 2, \ldots, N$ yang mempunyai $g_i = y$.

Selanjutnya mengambil $g^*$ yang mengindikasikan sebarang dari $\binom{N}{n}$ vektor – vektor yang mungkin dari z sejumlah n dan y sejumlah m, dan menyatakan replikasi permutasi dari $\hat{\theta}$,

$$\hat{\theta}^* = \hat{\theta}(g^*) = S(g^*, v) \qquad \ldots\ldots \quad (3.16)$$

Secara keseluruhan terdapat $\binom{N}{n}$ buah replikasi permutasi $\hat{\theta}^*$. Suatu distribusi yang meletakkan probabilitas $1/\binom{N}{n}$ pada masing-masing replikasi tersebut disebut sebagai distribusi permutasi dari $\hat{\theta}$, atau distribusi $\hat{\theta}^*$. Permutasi ASL dinyatakan sebagai probabilitas permutasi dari $\hat{\theta}^*$ yang melampaui $\hat{\theta}$,

$$ASL_{perm} = Prob_{perm}\{\ \hat{\theta}^* \geq \hat{\theta}\ \}$$

$$= \#\{\hat{\theta}^* \geq \hat{\theta}\ \} / \binom{N}{n} \qquad \ldots\ldots \quad (3.17)$$

Pernyataan $ASL_{perm}$ dalam (3.17) adalah identik akibat dari Lema Permutasi. Dengan kata lain penghitungan $ASL_{perm}$ bisa dilakukan dengan menghitung banyaknya replikasi permutasi dari $\hat{\theta}$ ( $\hat{\theta}^*$ pada (3.16) ) yang memenuhi syarat. Yaitu yang lebih besar atau sama dengan $\hat{\theta}$ pada (3.15). Kemudian banyaknya replikasi permutasi yang memenuhi syarat di atas dibagi dengan jumlah keseluruhan replikasi permutasi yang terjadi , yaitu sebanyak dari $\binom{N}{n}$ vektor – vektor yang mungkin dari z sejumlah n dan y sejumlah m.

Setelah penentuan **ASL** dan penghitungan statistik uji maka bila peluang yang telah ditentukan tidak dapat menjelaskan hasil-hasil itu, maka disimpulkan bahwa hipotesis nol tidak benar dan oleh karena itu ditolak. Bila telah ditentukan nilai tabel dari kriteria uji, maka $H_0$ ditolak bila nilai contoh dari kriteria ini lebih ekstrem daripada nilai tabel , selain itu $H_0$ diterima.

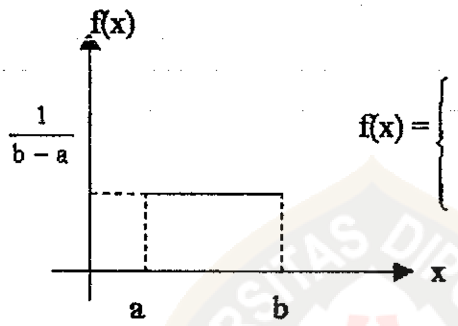
Uji permutasi didasarkan pada penyajian statistik order dari suatu data x = ( z, y ) dari suatu permasalahan dua sampel.

Selanjutnya dilakukan penghitungan $\mathbf{ASL}_{perm}$ dengan melakukan replikasi permutasi. Dalam praktek , $\mathbf{ASL}_{perm}$ biasanya didekati dengan Metode **Monte Carlo**, sesuai dengan Algoritma 3.1.

Replikasi dilakukan dengan mengambil hasil permutasi ( 2 himpunan ganda ) sebelumnya sebanyak B secara random. Metode **Monte Carlo** merupakan prosedur komputasi numerik yang melibatkan pengambilan sampel eksperimental dengan bilangan random. Penerapan metode **Monte Carlo** dalam replikasi permutasi untuk uji hipotesis ini dilakukan pada pengambilan sampel replikasi[ i ], ( i = 0, 1, 2, …, n). Suatu variabel berdistribusi uniform berarti bahwa probabilitas suatu variabel x yang berada pada suatu interval tertentu sebanding dengan rasio antara ukuran interval dengan jangkauannya.

Hal ini berarti setiap titik dalam jangkauan tersebut mempunyai kemungkinan yang sama untuk dipilih. Misalkan jangkauan nilai-nilai yang mungkin adalah dari a ke b untuk (b > a) maka probabilitas x berada pada interval Δx adalah Δx/(b-a). Bila digambar dalam grafik, fungsi kepadatan peluang merupakan sebuah garis lurus dengan tinggi 1/ (B – A) antara titik A dan B seperti yang ditunjukkan dalam gambar 3.4.

Gambar 3.4.




$$f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0 & , x \text{ di luarnya} \end{cases}$$

Untuk mendapatkan suatu urutan bilangan random pada interval [a,b], maka diperlukan suatu fungsi yang dapat mentransformasikan ke suatu interval tertentu. Formula transformasi yang banyak dipakai dalam pembangkitan bilangan random adalah *Linear Congruential Generators (LCGs).*

Sehingga akan didapat suatu barisan integer $z_1$, $z_2$, $z_3$, … yang didefinisikan dengan formula rekursiv sebagai berikut :

$$z_i = (a\, z_{i-1} + c) \pmod m$$

dengan a, c, m = konstanta > 0

$z_0$ = nilai awal.

Persamaan tersebut menyatakan bahwa untuk memperoleh $z_i$ dilakukan pembagian terhadap ($a\, z_{i-1}$ + c) dengan m dan $z_i$ merupakan sisa dari pembagian ini. Oleh karena itu $0 \le z_i \le m - 1$, dan untuk memperoleh bilangan acak yang diinginkan yaitu $U_i$ (untuk i = 1, 2, 3, …) pada [0,1], diambil $U_i = z_i / m$. Selain itu, integer m, a, c dan $z_0$ harus memenuhi $0 < m$, $a < m$, $c < m$ dan $z_0 < m$.

Penggunaan Metode **Monte Carlo** dalam penghitungan ini adalah pada pengambilan sampel eksperimental yaitu pengambilan vektor $g^*$ (1), $g^*$ (2), ..., $g^*$ (B), yang diambil secara random. Vektor-vektor yang akan diambil sebelumnya berarti mempunyai distribusi uniform. Untuk pembangkitan bilangan randomnya dapat dipilih suatu formula transformasi yang akan digunakan.

**Algoritma 3.1.**

Dasar Algoritma dari Uji Permutasi ASL

1. Melakukan pengurutan dan penggabungan dua buah sampel dan melakukan permutasi himpunan ganda dari gabungan sampel tersebut sehingga menghasilkan sejumlah vektor independen $g^*$.
2. Memilih sebanyak B vektor independen $g^*$ (1), $g^*$ (2), ..., $g^*$ (B), yang masing-masing terdiri dari z sejumlah n dan y sejumlah m, yang masing-masing dipilih secara random dari himpunan seluruh $\binom{N}{n}$ vektor-vektor yang mungkin. ( B biasanya diambil paling sedikit 1000).
3. Mengevaluasi replikasi permutasi dari $\hat{\theta}$ berhubungan dengan masing-masing vektor permutasi,

   $$\hat{\theta}^*(b) = S(g^*(b), v), \qquad b = 1,2,\ldots,B. \qquad \ldots.. \quad (3.18)$$

4. Melakukan aproksimasi $\mathbf{ASL}_{perm}$ dengan

   $$\widehat{\mathbf{ASL}}_{perm} = \#\{\hat{\theta}^*(b) \geq \hat{\theta}\}/B. \qquad \ldots.. \quad (3.19)$$

### 3.3 Uji ASL pada Bootstrap

Suatu uji hipotesis Bootstrap, didasarkan pada suatu statistik uji. Di sini $\hat{\theta}$ dinotasikan dengan t(x). Sehingga nilai achieved significance level-nya adalah

$$ASL = Prob_{H_0}\{ t(x^*) \geq t(x) \} \quad ..... \quad (3.20)$$

t (x) = nilai yang diamati

$x^*$ = variabel random dari distribusi hipotesis nol. ( Sebut distribusi $F_0$ ).

Dalam uji hipotesis Bootstrap , digunakan metode "plug-in" dalam melakukan estimasi $F_0$. Dengan kata lain dalam mengestimasi $F_0$ digunakan fungsi yang sama dengan distribusi empiris pada $\hat{F}_0$ . Selanjutnya memberi notasi x pada sampel yang telah digabungkan dan mengambil distribusi empirisnya sebagai $\hat{F}_0$ , meletakkan probabilitas 1 / (n+m) pada masing-masing anggota x. Algoritma 3.2 menunjukkan bagaimana $ASL_{boot}$ dihitung.

**Algoritma 3.2.**

<u>Dasar Algoritma dari Statistik Uji Bootstrap untuk Pengujian F = G</u>

1. Memilih sebanyak B sampel ukuran n + m dengan pengembalian dari x. Sebut n observasi pertama sebagai $z^{*b}$ dan m observasi selanjutnya $y^{*b}$, untuk b = 1, 2, ..., B.
2. Mengevaluasi t (.) pada masing-masing sampel,

$$t(x^{*b}) = \bar{z}^{*b} - \bar{y}^{*b}, \qquad b = 1, 2, \ldots, B. \quad ..... \quad (3.21)$$

3. Melakukan aproksimasi $ASL_{boot}$ dengan

$$\widehat{ASL}_{boot} = \#\{ t(x^{*b}) \geq t_{obs} \} / B. \quad ..... \quad (3.22)$$

$t_{obs} = t(x)$ merupakan nilai observasi dari statistiknya.

Pengujian yang lebih akurat didapatkan dengan menggunakan statistik dari Student. Sehingga pada uji di atas, selain t (x) = $\bar{z} - \bar{y}$ digunakan

$$t(x) = \frac{\bar{z} - \bar{y}}{\bar{\sigma}\sqrt{1/n + 1/m}} \qquad \ldots\ldots \quad (3.23)$$

Algoritma 3.2 menguji hipotesis nol yang menyatakan sifat identik pada dua buah populasi, yaitu F = G. Jika yang diinginkan hanya pengujian apakah nilai rata-rata mereka sama, dapat digunakan suatu pendekatan dengan statistik t dua sampel ( 3.23). Di bawah hipotesis nol dan mengasumsikan populasi normal dengan variansi yang sama, maka memiliki distribusi Student -t dengan n + m − 2 derajat kebebasan. Seandainya tidak diasumsikan nilai variansi yang sama pada dua buah populasi , maka dapat digunakan

$$t(x) = \frac{\bar{z} - \bar{y}}{\sqrt{\bar{\sigma}_1^2/n + \bar{\sigma}_2^2/m}} \qquad \ldots\ldots \quad (3.24)$$

dengan $\bar{\sigma}_1^2 = \sum_1^n (z_i - \bar{z})^2/(n-1)$ , $\bar{\sigma}_2^2 = \sum_1^m (y_i - \bar{y})^2/(m-1)$ .

Dalam uji hipotesis dengan bootstrap untuk membandingkan dua buah rata-rata, tidak ada alasan yang memaksa mengasumsikan variansi yang sama sehingga asumsi tersebut dapat tidak dipergunakan.

Sehingga untuk meneruskan perhitungannya dibutuhkan estimasi dari F dan G yang hanya menggunakan asumsi dari sebuah rata-rata yang biasa. Dengan mengambil $\bar{x}$ sebagai rata-rata sampel yang telah dikombinasikan, maka kedua buah sampel dapat dirubah sehingga keduanya memiliki rata-rata $\bar{x}$, dan kemudian melakukan resampel tiap populasinya secara terpisah. Prosedurnya ditunjukkan pada perincian Algoritma 3.3.