BAB II

TINJAUAN PUSTAKA

2.1 Pendahuluan

Untuk melakukan pembahasan mengenai materi di skripsi ini, diperlukan teori-teori yang mendukung. Pada bab ini akan diuraikan beberapa teori yang mendukung penulisan skripsi. Uraian dimulai dengan membahas statistika parametrik dan nonparametrik, pengujian hipotesis, *One-Way Layout,* dan pengujian untuk lebih dari dua sampel saling bebas.

2.2. Statistika Parametrik dan Statistika Nonparametrik

Pengantar statistika umumnya menekankan pembahasan pada prosedur statistika parametrik. Statistika parametrik digunakan untuk menganalisis data yang berskala interval atau rasio yang diambil dari populasi yang berdistribusi normal. Untuk dapat menggunakan statistika inferensial ini, ada beberapa syarat yang harus dipenuhi, yaitu data harus berdistribusi normal dan sampel yang diteliti harus diambil secara acak. Salah satu karakteristik prosedur ini adalah bahwa kelayakan penggunaanya untuk maksud inferensi (pengujian atau penaksiran) bergantung pada asumsi tertentu. Prosedur inferensial dalam analisis varians misalnya, mengasumsikan bahwa sampel telah ditarik dari populasi yang berdistribusi normal dengan varians yang sama.

Karena populasi yang kita kaji tidak selalu memenuhi asumsi yang mendasari uji parametrik, kita kerap kali membutuhkan prosedur inferensial dengan kesahihan (*validity*) yang tidak bergantung pada asumsi tertentu. Dalam mengatasi masalah tersebut, prosedur statistika nonparametrik memenuhi kebutuhan ini karena tetap sahih meski hanya berlandaskan asumsi yang sangat umum (Daniel, 1989).

Uji statistika nonparametrik adalah uji yang modelnya tidak menetapkan syarat mengenai parameter populasi yang merupakan induk sampel penelitiannya (Siegel, 1997). Statistika nonparametrik digunakan untuk menganalisis data yang sekurang-kurangnya berskala ordinal dari populasi yang bebas distribusi (tidak harus berdistribusi normal). Dua tipe utama prosedur statistik yang dianggap nonparametrik adalah: 1) prosedur-prosedur nonparametrik murni, 2) prosedur-prosedur bebas-distribusi (*distribution-free procedures*). Statistika nonparametrik dapat digunakan untuk menguji hipotesis yang menggunakan satu sampel dan lebih dari dua sampel. Dalam uji satu sampel kita menarik suatu sampel random dan kemudian menguji hipotesis bahwa sampel ini ditarik dari suatu populasi dengan populasi tertentu, sedangkan dalam kasus lebih dari dua sampel kita menguji hipotesis yang menyatakan bahwa beberapa sampel telah ditarik dari populasi yang sama atau dari populasi dengan parameter yang sama.

Keunggulan atau kelebihan statistik nonparametrik adalah (Daniel, 1989):

1. Karena kebanyakan prosedur nonparametrik memerlukan asumsi dalam jumlah yang minimum, maka kemungkinan terjadinya kesalahan pun kecil.
2. Untuk beberapa prosedur nonparametrik, perhitungan dapat dilakukan dengan cepat dan mudah, terutama bila terpaksa dikerjakan secara manual. Jadi penggunaan prosedur ini lebih menghemat waktu dalam perhitungan.
3. Para peneliti dengan dasar matematika serta statistika yang kurang biasanya menganggap bahwa konsep dan metode prosedur nonparametrik mudah dipahami.

4. Jika sampelnya kecil statistik nonparametrik dapat digunakan, kecuali kalau sifat distribusi populasinya diketahui secara pasti.

Sedangkan kekurangan atau kelemahan statistik nonparametrik diantaranya:

1. Karena perhitungan yang dibutuhkan untuk kebanyakan prosedur nonparametrik cepat dan sederhana, prosedur ini kadang-kadang digunakan untuk kasus yang lebih tepat bila ditangani dengan prosedur parametrik. Cara seperti ini sering menyebabkan pemborosan informasi.
2. Kendatipun prosedur nonparametrik terkenal karena prinsip perhitungannya yang sederhana, pekerjaan hitung menghitung (*arithmetic*)-nya sendiri sering membutuhkan banyak tenaga serta menjemukan.

Setelah melihat kekurangan dan kelebihan statistik nonparametrik, dibawah ini akan dikemukakan beberapa situasi yang tepat bila ditangani dengan prosedur nonparametrik:

1. Bila hipotesis yang harus diuji tidak melibatkan suatu parameter populasi.
2. Bila data diukur menggunakan skala yang lebih tepat menggunakan prosedur statistika nonparametrik. Sebagai contoh, data mungkin terdiri atas data hitung atau data peringkat, sehingga menghalangi penerapan prosedur parametrik yang semestinya lebih tepat.
3. Bila asumsi yang diperlukan tidak terpenuhi dengan menggunakan prosedur parametrik. Sebagai contoh, dalam rancangan riset percobaan suatu proyek mungkin menganjurkan penggunaan prosedur parametrik tertentu. Tetapi dalam pemeriksaan data mungkin salah satu atau beberapa asumsi yang mendasari pengujian tersebut tidak terpenuhi, sehingga prosedur statistik nonparametrik bisa menjadi solusinya.

2.3 Pengujian Hipotesis

Langkah pertama dalam prosedur pengujian hipotesis adalah menyatakan hipotesis nol-nya (H_0). Hipotesis nol ini adalah suatu hipotesis tentang ada tidaknya perbedaan atau pengaruh dari parameter yang diuji. Hipotesis ini pada umumnya diformulasikan untuk ditolak. Apabila ditolak, maka hipotesis pengganti (H_1) dapat diterima. Hipotesis pengganti ini merupakan hipotesis penelitian dari si pembuat eksperimen, yang dinyatakan secara operasional. Hipotesis penelitian adalah prediksi yang diturunkan dari teori yang sedang diuji (Siegel, 1997).

Bila kita hendak membuat keputusan mengenai perbedaan, kita menguji H_0 terhadap H_1. H_1 merupakan pernyataan yang diterima jika H_0 ditolak. Uji hipotesis bisa dua sisi (*two-sides/two-tailed/nondirectional*;tanpa arah/dwi arah), bisa juga satu sisi (*one-sides/one-tailed/directional*; searah/eka arah). Yang berikut ini adalah contoh pernyataan hipotesis nol dan hipotesis tandingannya bila parameter yang ingin kita ketahui adalah rata-rata populasi μ_1 untuk populasi 1, dan rata-rata populasi μ_2 untuk populasi 2, dengan perumusan hipotesisnya:

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2 \qquad (2.1)$$

Disini hipotesis nol menyatakan bahwa rata-rata kedua populasi itu sama. Sedangkan hipotesis tandingannya menyatakan bahwa rata-rata keduannya tidak sama. Untuk uji satu sisi atau uji eka arah perumusan hipotesis statistiknya adalah:

$H_0 : \mu_1 \geq \mu_2$ vs $H_1 : \mu_1 < \mu_2$ atau

$H_0 : \mu_1 \leq \mu_2$ vs $H_1 : \mu_1 > \mu_2$

Guna menguji hipotesis itu, peneliti memilih statistik uji (*test statistic*) yang paling tepat apakah Z, atau t (Daniel, 1989).

Pengujian untuk data yang lebih dari dua sampel adalah perluasan uji median, uji Kruskal-Wallis, dan perbandingan berganda. Ketiga metode tersebut digunakan untuk menguji populasi yang mempunyai median sama atau tidak. Selain ketiga metode tersebut terdapat uji Jonckheere Terpstra untuk pengujian lebih dari dua sampel, namun uji ini dikhususkan untuk menguji data yang mempunyai median berurutan (*ordered*). Uji Jonckheere Terpstra mempunyai tiga model dengan median meningkat, menurun, atau berbentuk seperti payung. Sebelum membahas keempat metode tersebut, disini akan dibahas terlebih dahulu salah satu desain eksperimental *one-way layout* yang berhubungan dengan pengujian statistik.

2.4 *One-Way Layout*

One-way adalah desain eksperimental yang umum digunakan dalam penelitian ilmiah. Dengan desain eksperimental ini dapat memudahkan menguji hipotesis statistik yang mempunyai kelompok kontrol dan kelompok perlakuan. Misalnya, tingkat pengobatan dalam penelitian terhadap pemberian peningkatan dosis, apakah respon akan menurun atau meningkat dengan meningkatkan tingkat dosis. Selain itu, penelitian sederhana lainnya dalam *one-way* misalnya, mata pelajaran sebagai unit eksperimental diberikan kepada dua atau lebih kelompok perlakuan yang telah ditetapkan. Setelah setiap kelompok diberikan perlakuan, kemudian dilakukan pengukuran (sering disebut respon perlakuan) (Juneau, 2006). Secara umum X_{ij} = Unit ke-j ($1 \leq j \leq n_i$) dengan perlakuan ke-i ($1 \leq i \leq k$)

Perlakuan 1: terdapat n_1 unit percobaan dengan perlakuan 1: $x_{11}, x_{12}, \ldots, x_{1n_1}$

Perlakuan 2: terdapat n_2 unit percobaan dengan perlakuan 2: $x_{21}, x_{22}, \ldots, x_{2n_2}$

Perlakuan i: terdapat n_i unit percobaan dengan perlakuan i: $x_{i1}, x_{i2}, \ldots, x_{in_i}$

Perlakuan k: terdapat n_k unit percobaan dengan perlakuan k: $x_{k1}, x_{k2}, \ldots, x_{kn_k}$

Data terdiri dari N = $\sum_{i=1}^{k} n_i$ yang jika disusun dalam tabel adalah sebagai berikut:

Tabel 2.1 Struktur Data *k* Sampel Perlakuan

Perlakuan					
1	2	…	i	…	k
x_{11}	x_{21}	…	x_{i1}	…	x_{k1}
x_{12}	x_{22}	…	x_{i2}	…	x_{k2}
⋮	⋮	⋮	⋮	⋮	⋮
X_{1n_1}	X_{2n_2}	…	X_{in_i}	…	X_{kn_k}

Asumsi:

1. $X_1, X_2, \ldots, X_k$ merupakan sampel acak yang saling *independen* (bebas).
2. Untuk setiap *i* (1, 2, ..., k), n_i sampel acak ($X_1, X_2, \ldots, X_k$) berdistribusi kontinu dengan fungsi distribusi F_i.
3. Fungsi distribusi $F_1, \ldots, F_k$ adalah sama, tetapi ada perbedaan dalam hal lokasi. Sehingga fungsi distribusinya terhubung melalui hubungan:

   $F_i(x) = F(x - \tau_i), -\infty < x < \infty$ (2.2)

   Untuk i = 1, ..., k, di mana F adalah fungsi distribusi kontinu dengan median θ

dan τ_i adalah efek perlakuan ke-i. Semua asumsi tersebut berhubungan langsung dengan *one-way layout* yang umumnya terkait dengan asumsi normal yaitu, semua asumsi tersebut setara dengan representasi $X_{ij} = \theta + \tau_i + \varepsilon_{ij}$, i = 1, ..., k, j = 1, ..., n_i, di mana θ adalah median setiap kelompok, τ_i adalah efek perlakuan i, dan ε yang membentuk sampel acak dari distribusi kontinu dengan median θ.

Hipotesis untuk *one-way layout* adalah:

H_0: τ_1, =... = τ_k (2.3)

Hipotesis nol ini menegaskan bahwa setiap distribusi yang mendasari F_1, ..., F_k adalah sama, sesuai dengan $F_1 = F_2 = ... = F_k$ melawan semua bentuk τ_i berbeda dari H_0 (Hollander dan wolfe, 1999).

2.5 Pengujian Untuk Lebih Dari Dua Sampel Bebas

Dalam melakukan pengujian untuk data sampel yang lebih dari dua sampel, metode yang dapat digunakan adalah Perluasan Uji Median, uji Kruskal-Wallis, uji Perbandingan Berganda, dan uji Jonckheere Terpstra. Metode-metode tersebut digunakan untuk menguji populasi yang mempunyai median sama atau tidak, terkecuali uji Jonckheere Terpstra. Uji Jonckheere Terpstra ini dikhususkan untuk menguji data yang mempunyai median berurutan (*ordered*).

2.5.1. Perluasan Uji Median

Perluasan uji median mempunyai asumsi (Karyana dkk., 2011):

- Masing-masing sampel adalah sampel acak berukuran n_i yang ditarik dari salah satu diantara k populasi yang diamati.
- Pengamatan dilakukan secara bebas baik di dalam maupun diantara sampel.
- Skala pengukuran minimal ordinal.
- Jika semua populasi memiliki median sama, maka untuk masing-masing populasi peluang p suatu nilai pengamatan lebih besar dari median gabungan sama besar.

Hipotesis untuk uji ini adalah sebagai berikut:

H_0 : semua k populasi mempunyai median yang sama

H_1 :paling sedikit satu populasi mempunyai median berbeda dengan yang lain.

Statistik ujinya adalah sebagai berikut:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{k}\left[\frac{(O_{ij} - E_{ij})^2}{E_{ij}}\right] \tag{2.4}$$

$$E_{ij} = \left(\frac{n_i . n_{.j}}{N}\right) \tag{2.5}$$

Dimana i = 1, 2, ...,k dan j = 1, 2, ...,n_i

Dengan kriteria ujinya:

Tolak H_0 jika $\chi^2 > \chi^2{}_{(1-\alpha)(k-1)(r-1)}$

2.5.2. Uji Kruskal-Wallis

Untuk menentukan apakah k sampel independen berasal dari populasi yang berbeda, analisis varian ranking satu arah Kruskal-Wallis menitik beratkan pada *ranking* dari data yang digunakan (Siegel, 1997). Uji Kruskal-Wallis menguji hipotesis-nol bahwa k sampel berasal dari populasi yang sama atau populasi identik, dalam hal rata-ratanya. Uji ini membuat anggapan bahwa variabel yang diamati mempunyai distribusi kontinu. Uji ini menuntut pengukuran variabelnya paling tidak dalam skala ordinal.

Asumsi-asumsi yang terdapat pada uji Kruskal-Wallis adalah sebagai berikut (Daniel, 1989):

1. Data untuk analisis terdiri atas k sampel acak berukuran n_1, n_2, ..., n_k.
2. Pengamatan bisa dilakukan baik di dalam maupun di antara sampel-sampel.
3. Variabel yang diteliti kontinu.
4. Skala pengukuran yang digunakan setidaknya ordinal.
5. Populasi-populasi identik kecuali dalam hal lokasi yang mungkin berbeda untuk sekurang-kurangnya satu populasi.

Hipotesis untuk uji Kruskal-Wallis adalah:

H_0: tidak ada perbedaan nilai median populasi ($\theta_1 = \theta_2 = \cdots = \theta_k$)

H_1: minimal ada satu pasang median populasi yang tidak sama $\left(\theta_i \neq \theta_j\right)$

Statistik uji Kruskal-Wallis ini, masing-masing N observasinya digantikan dengan *ranking*nya. Semua nilai dalam seluruh k sampel yang digunakan, diurutkan (*ranking*) dalam satu rangkaian. Nilai yang terkecil digantikan dengan ranking 1, setingkat di atas yang terkecil digantikan dengan ranking 2, dan yang terbesar dengan *ranking N*. Dimana *N* adalah jumlah seluruh observasi independen dalam *k* sampel tersebut. Setelah semua nilai dalam *k* sampel yang digunakan diurutkan (*ranking*) hitung jumlah *ranking* dalam masing-masing sampel. Uji Kruskal-Wallis menentukan apakah jumlah *ranking* itu berlainan atau berbeda, sehingga sangat kecil kemungkinan bahwa semua sampel tersebut ditarik dari populasi yang sama. Jika seluruh *k* sampel tersebut memang benar-benar dari populasi yang sama atau populasi yang identik, yakni jika H_0 benar, maka *H* (statistik yang dipergunakan dalam uji Kruskal-Wallis ini dan didefinisikan dengan rumus di bawah ini) berdistribusi chi-kuadrat dengan *db* = *k-1*, dengan syarat bahwa ukuran *k* sampel itu tidak terlalu kecil. Statistik uji *H* adalah:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{{R_i}^2}{n_i} - 3(N+1) \qquad (2.6)$$

Dimana:

k = banyak sampel
n_i = banyak unit dalam sampel ke-i
N = $\sum n_i$ = banyak unit dalam semua sampel
R_i = jumlah peringkat dalam sampel ke-i.

Kriteria uji atau kaidah pengambilan keputusan uji Kruskal-Wallis dilakukan berdasarkan hipotesis nol. Jika terdapat lebih dari lima unit dalam setiap kelompok, yakni n > 5, maka nilai statistik *H* dapat ditentukan dengan menggunakan Tabel nilai kritis uji Kruskal-Wallis untuk mengetahui taraf nyata. Jika *H* lebih besar atau sama dengan chi-kuadrat yang ditunjukkan dalam Tabel nilai kritis chi-kuadrat untuk tingkat signifikansi yang telah ditetapkan, dengan $db = k - 1$, maka H_0 ditolak pada tingkat signifikansi tersebut.

Jika banyaknya unit dalam masing-masing sampelnya kurang atau sama dengan lima n ≤ 5, maka pendekatan chi-kuadrat pada distribusi sampling *H* tidak cukup baik. Nilai-nilai kemungkinan tersebut disajikan dalam Tabel nilai kritis uji Kruskal-Wallis. Jika terjadi angka sama antara dua nilai atau lebih, setiap nilai mendapatkan *ranking* yang sama, yaitu rata-rata *ranking*nya.

Karena angka sama mempengaruhi nilai statistik *H*, perlu ditambahkan faktor koreksi Φ_1 untuk menghitung statistik uji. Untuk menambahkan faktor koreksi karena angka sama tersebut, *H* dihitung dengan rumus (2.6) dan kemudian dibagi dengan:

$$\Phi_1 = 1 - \frac{\sum T}{N^3 - N} \tag{2.7}$$

Dimana:

$T = t^2 - 1$ (jika t adalah banyak unit-unit berangka sama dalam serangkaian skor berangka sma).

N = banyak observasi dalam seluruh k sampel bersama-sama, yakni N = $\sum n_i$.

Dengan demikian, rumus umum untuk *H* yang telah dikoreksi karena adanya angka sama adalah:

$$H_k = \frac{\frac{12}{N\,(N+1)} \sum_{i=1}^{k} \frac{{R_i}^2}{n_i} - 3\,(N+1)}{1 - \frac{\sum T}{N^3 - N}} \tag{2.8}$$

Dengan adanya faktor koreksi untuk angka sama tersebut, hasil nilai statistik H_k akan lebih signifikan bila dibandingkan tanpa faktor koreksi. Oleh karena itu, jika kita dapat menolak H_0 menggunakan faktor koreksi tersebut, kita akan memperoleh tingkat signifikansi yang lebih meyakinkan dibandingkan tanpa menggunakan koreksi (Siegel, 1997).

2.5.3. Uji Perbandingan Berganda

Apabila hasil dengan uji Kruskal-Wallis menyatakan penolakan terhadap H_0, maka proses selanjutnya melakukan perbandingan berganda (Karyana dkk., 2011). Rumus untuk melakukan uji perbandingan berganda adalah:

1. Jika ukuran sampel berbeda

$$\left|\bar{R}_i - \bar{R}_j\right| \leq Z_{(1-\{\frac{\alpha}{k(k-1)}\}}\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)} \tag{2.9}$$

Dimana $N = n_1 + n_2 + \ldots + n_k$ atau N merupakan banyaknya hasil pengamatan dari semua sampel yang digabungkan. Apabila rumus diatas terpenuhi, maka kesimpulannya R_i dan R_j tidak berbeda.

2. Jika ukuran sampel sama

$$\left|\bar{R}_i - \bar{R}_j\right| \leq Z_{(1-\{\frac{\alpha}{k(k-1)}\}}\sqrt{k(N+1)/6} \tag{2.10}$$

Apabila persamaan diatas terpenuhi, maka kesimpulannya R_i dan R_j tidak berbeda.

3. Jika ukuran sampel berbeda tetapi terdapat angka sama

$$\left|\overline{R_i} - \bar{R}_j\right| \leq Z_{(1-\{\frac{\alpha}{k(k-1)}\})}\sqrt{\frac{[N(N^2-1)-(\sum t^3-\sum t)]\left[\frac{1}{n_i}-\frac{1}{n_j}\right]}{12(N-1)}} \tag{2.11}$$

Apabila persamaan diatas terpenuhi, maka kesimpulannya R_i dan R_j tidak berbeda.

4. Jika ukuran sampel sama tetapi terdapat angka yang sama

$$\left|\overline{R_i} - \bar{R}_j\right| \leq Z_{(1-\{\frac{\alpha}{k(k-1)}\})}\sqrt{\frac{k[N(N^2-1)-(\sum t^3-\sum t)]}{6N(N-1)}} \tag{2.12}$$

Apabila persamaan diatas terpenuhi, maka kesimpulannya R_i dan R_j tidak berbeda.

2.5.4. Uji Jonckheere Terpstra

Pengujian hipotesis alternatif berurutan untuk sampel yang saling bebas ditemukan oleh Aimable Robert Jonckheere dan Terpstra pada tahun 1954, sehingga disebut uji Jonckheere Terpstra. Uji Jonckheere Terpstra digunakan untuk mengetahui apakah median populasi sama berdasarkan k sampel yang saling bebas atau mempunyai kecenderungan meningkat, menurun atau berbentuk seperti payung, berdasarkan statistik peringkat median (Jonckheere, 1954).

Tabel 2.2 Struktur Data k Sampel Perlakuan

Perlakuan					
1	2	⋯	i	⋯	k
x_{11}	x_{21}	...	x_{i1}	...	x_{k1}
x_{12}	x_{22}	...	x_{i2}	...	x_{k2}
⋮	⋮	⋮	⋮	⋮	⋮
X_{1n_1}	X_{2n_2}	...	X_{in_i}	...	X_{kn_k}

Dimana:
X_{ij} = unit ke-j dari perlakuan ke-i (i = 1, ..., k); j (1, ..., n_i)
n_i = banyakya unitpada perlakuan ke-i
k = banyaknya perlakuan (i= 1, 2, ..., k).
Statistik uji Jonkheere Terpstradapat ditulis sebagai berikut:

$$J = \sum_{i=1}^{l-1} \sum_{l=2}^{k} U_{\tau_i \tau_l} \quad (2.13)$$

$$J = \sum_{i=1}^{l-1} (U_{\tau_i \tau_2} + U_{\tau_i \tau_3} + \cdots + U_{\tau_i \tau_k})$$

$$= U_{\tau_1 \tau_2} + U_{\tau_1 \tau_3} + \cdots + U_{\tau_1 \tau_k}$$

$$U_{\tau_i \tau_l} = \sum U_{\tau_i \tau_l}$$

$$U_{\tau_i\tau_l} = \begin{cases} 1\ ;\ x_i < x_l \\ 0\ ;\ x_i > x_l \\ \frac{1}{2}\ ;\ x_i = x_l \end{cases}$$

Dengan $U_{\tau_i\tau_l}$ adalah banyaknya pasangan hasil pengamatan yang dalam hal ini τ_i lebih kecil dari τ_l. Jadi, kita memperbandingkan hasil-hasil pengamatan dalam semua pasangan sampel. Kita memperbandingkan masing-masing nilai pengamatan dalam sampel pertama dengan setiap nilai pengamatan dalam sampel kedua, dan apabila nilai pengamatan dari sampel pertama lebih kecil daripada nilai pengamatan di sampel kedua, kita memberikan skor 1 bagi pasangan yang bersangkutan. Dan apabila nilai pengamatan dari sampel pertama lebih besar daripada nilai pengamatan di sampel kedua, maka skor yang kita berikan bagi pasangan tersebut adalah 0. Jika terjadi angka sama dalam menghitung $U_{\tau_i\tau_l}$ berilah skor ½ untuk setiap unit dimana $\tau_i = \tau_l$. Dengan kata lain, berilah skor ½ setiap kali menjumpai angka sama ketika sedang memperbandingkan nilai-nilai pengamatan (Daniel, 1989).

Asumsi-asumsi yang terdapat pada uji Jonckheere Terpstra adalah sebagai berikut:

1. Data untuk analisis terdiri atas k sampel acak berukuran n_1, n_2, ..., n_k, yang berturut-turut berasal dari populasi-populasi *1, 2, ...,k*.
2. Nilai-nilai pengamatan tidak berkaitan baik di dalam maupun di antara sampel-sampel.
3. Variabel yang diteliti kontinu.
4. Skala pengukuran sekurang-kurangnya ordinal.
5. Populasi asal sampel identik kecuali adanya kemungkinan perbedaan dalam parameter-parameter lokasi.

Statistik J digunakan untuk menguji hipotesis nol bahwa tidak ada perbedaan median antar k sampel, melawan hipotesis alternatif antar k sampelyang memiliki median berurutan. Hipotesis dapat ditulis dengan:

H_0: tidak ada perbedaan nilai median populasi ($\theta_1 = \theta_2 = ... = \theta_k$)

H_1: median kelompok satu lebih kecil dari kelompok dua, kelompok dua, ..., kelompok k – 1 lebih kecil dari kelompok k ($\theta_1 < \theta_2 < ... < \theta_k$) (2.14)

Nilai-nilai dalam tabel statistik J menunjukan peluang yang berhubungan dengan nilai statistik uji J apakah sama besar atau lebih besar daripada nilai J pada tabel sesuai dengan ukuran sampel dan taraf nyata (α). Jika nilai $J \geq J_\alpha$ maka tolak H_0 (Hollander dan Wolfe, 1999). Karena distribusi J memiliki kesimetrisan yang tertentu, maka kita boleh mendapatkan nilai-nilai kritis untuk konfigurasi-konfigurasi yang tidak dalam urutan demikian dengan mengatur kembali ketiga ukuran sampel sehingga memiliki urutan ukuran yang meningkat sebelum kita mengacu ke tabel. Sebagai contoh, kalau kita ingin mendapatkan nilai-nilai kritis untuk ukuran sampel $n_1 = 5$, $n_2 = 7$, $n_3 = 3$ kita boleh mengacu ke Tabel pada $n_1 = 3$, $n_2 = 5$, $n_3 = 7$.

Aproksimasi atau pendekatan untuk sampel besar digunakan jika ukuran sampel lebih dari 8 (n > 8). Jika sampel berukuran kurang dari atau sama dengan 8 (n ≤ 8) maka gunakan rumus untuk sampel kecil. J mengikuti distribusi normal dengan rata-rata 0 dan varians 1. Rumus untuk menghitung nilai ekspektasi dan varians J adalah:

$$E_0(J) = N^2 - \sum_{i=1}^{k} {n_i}^2/4 \quad (2.15)$$

$$Var_0(J) = \frac{N^2(2N+3) - \sum_{i=1}^{k} {n_i}^2(n_i+3)}{72} \quad (2.16)$$

Bentuk standar dari rumus sampel besar adalah:

$$J_{sb} = \frac{J - E_0(J)}{\sqrt{Var_0(J)}} = \frac{J - \left[\left(N^2 - \sum_{i=1}^{k} n_i{}^2\right)\right]/4}{\sqrt{\left[N^2(2N+3) - \sum_{i=1}^{k} n_i{}^2(2n_i+3)\right]/72}} \qquad (2.17)$$

Kriteria uji untuk statistik J_{sb} adalah dengan menolak H_0 jika P_{value} dari $J_{sb} < \alpha$, dimana nilai J_{sb} dalam distribusi normal sesuai dengan taraf nyata yang digunakan. Ketika terjadi angka sama dalam aproksimasi sampel besar, faktor koreksi harus ditambahkan dalam perhitungan. Nilai variansi *J* menjadi sebagai berikut:

$$Var_0(J) = \Big\{\frac{1}{72}\big[N(N-1)(2N - \sum_{i=1}^{k} n_i(n_i-1)(2n_i+5) - \sum_{i=1}^{g} t_i(t_i-1)(2t_i+5)\big] + \frac{1}{36N(N-1)(N-2)}\big[\sum_{i=1}^{k} n_i(n_i-1)(n_i-2)\big]\big[\sum_{i=1}^{g} t_i(t_i-1)(t_i-2)\big] + \frac{1}{8N(N-1)}\big[\sum_{i=1}^{k} n_i(n_i-1)\big]\big[\sum_{i=1}^{g} t_i(t_i-1)\big]\Big\} \qquad (2.18)$$

Dimana *g* menunjukan banyaknya kelompok angka sama yang terdapat di dalam semua hasil pengamatan dan t_i adalah banyaknya anggota kelompok angka sama setiap perlakuan ke i. Akibat dari pengaruh angka sama di variansi *J*, modifikasi berikut diperlukan untuk menerapkan aproksimasi sampel besar ketika ada angka yang sama. Rumus untuk menghitung J_{sb} menggunakan modifikasi Mann-Whitney adalah:

$$J_{msb} = \frac{J\left[N^2 - \sum_{i=1}^{k} n_i{}^2\right]}{\{Var_0(J)\}^{1/2}} \qquad (2.19)$$