BAB III

DISTRIBUSI SAMPLING DAN UJI SIGNIFIKANSI r_s

Distribusi sampling adalah distribusi dari semua harga yang mungkin dijalani oleh suatu statistik, bila statistik itu dihitung dari sampel-sampel yang sama besarnya.

3.1 Distribusi Nol Eksak Dari r_s

Jika X dan Y adalah dua variabel yang saling bebas, dengan sampel bivariat $\{(x_1,y_1),(x_2,y_2),\ldots\ldots,(x_n,y_n)\}$. Karena X dan Y saling bebas, maka himpunan pasangan ranknya adalah $\{(r_1,s_1),(r_2,s_2),\ldots\ldots,(r_n,s_n)\}$. Rank tersebut merupakan salah satu diantara sejumlah n! buah himpunan pasangan rank sampel yang mungkin diperoleh dengan peluang yang sama besar. Oleh karena itu statistik

$$r_s = 1 - \frac{6\sum_{i=1}^{n}(r_i - s_i)^2}{n(n^2-1)}$$

adalah suatu statistik bebas distribusi (distribution-free statistik), dan distribusi sampling dari r_s dapat ditentukan kemudian.

Jika ruang sampel adalah himpunan semua pasangan rank sampel dari $\{(r_1,s_1),(r_2,s_2),\ldots\ldots\ldots,(r_n,s_n)\}$

yang diperoleh dengan jalan membuat permutasi pada $\{(r_1,s_1),(r_2,s_2),\ldots\ldots,(r_n,s_n)\}$, dimana $r_1,r_2,\ldots\ldots,r_n$ diambil tetap, dan $s_1,s_2,\ldots\ldots,s_n$ dipermutasikan, maka terdapat sebanyak n! buah titik sampel.

Jika U_r adalah banyaknya titik sampel, yaitu himpunan pasangan rank $\{(r_1,s_1),(r_2,s_2),\ldots\ldots,(r_n,s_n)\}$, yang memberikan harga statistik r_s sebesar r , maka distribusi nol dari r_s adalah :

$$f_{r_s}(r) = \frac{U_r}{n!} \quad \ldots\ldots\ldots\ldots\ldots\ldots(3.1.1)$$

dimana :

$f_{r_s}(r)$ = distribusi peluang dari r_s

U_r = banyaknya titik sampel yang memberikan harga r

n = banyaknya titik sampel.

Theorema 3.1.1.

Distribusi nol dari r_s simetri terhadap titik nol.

Bukti :

Pandang X dan Y dua variabel random yang saling bebas, dengan sampel random bivariat $\{(X_1,Y_1), (X_2,Y_2),\ldots\ldots,X_n,Y_n)\}$. Asumsikan kesemua n pasangan pengamatan $\{(x_1,y_1), (x_2,y_2),\ldots\ldots,(x_n,y_n)\}$ sudah disusun terlebih dahulu berdasarkan urutan besar relatif komponen X. Jadi kita dapatkan himpunan pasangan rank $\{(1,s_1),(2,s_2),\ldots\ldots(n,s_n)\}$.

Pandang variabel random D $=\sum_{i=1}^{n} D_i^2 = \sum_{i=1}^{n}(R_i-S_i)^2$ dimana R_i dan S_i masing-masing merupakan variabel rank untuk pengamatan x_i dan y_i, dan nilai dari variabel random D ini adalah d.

Maka didapatkan :

$$d_i = r_i - s_i$$
$$= i - s_i$$

Sehingga :

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (i - s_i)^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1.2)$$

Himpunan pasangan rank $\{(1,s_1),(2,s_2),\ldots\ldots,(n,s_n)\}$ ini terdapat himpunan pasangan rank sekawannya yaitu $\{(1,s_n),(2,s_{n-1}),\ldots\ldots,(n,s_1)\}$,

dimana :

$$d'_i = i - s_{n+1-i}$$
$$\text{untuk } i = 1,2,3,\ldots,n$$

Sehingga :

$$\sum_{i=1}^{n} d'_i = \sum_{i=1}^{n} (i-s_{n+1-i})$$
$$= (1-s_n)+(2-s_{n-1})+\ldots\ldots+ (n-s_1)$$
$$= (n-s_1)+(n-1-s_2)+(n-2-s_2)+\ldots\ldots+(2-s_{n-1})+(1-s_n)$$
$$= \sum_{i=1}^{n} (n+1-i-s_i) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1.3)$$

Sehingga didapatkan bahwa :

$$\sum_{i=1}^{n} (d_i+d'_i) = \sum_{i=1}^{n} (i-s_i + i-s_{n+1-i})$$

$$= \sum_{i=1}^{n}(i-s_i) + \sum_{i=1}^{n}(i-s_{n+1-i})$$

Dari persamaan (3.1.2) dan (3.1.3) maka didapatkan :

$$\sum_{i=1}^{n}(d_i+d'_i) = \sum_{i=1}^{n}(i-s_i+n+1-i-s_i)$$

$$= \sum_{i=1}^{n}(n+1-2s_i)$$

Sehingga :

$$\sum_{i=1}^{n}(d_i+d'_i)^2 = \sum_{i=1}^{n}(n+1-2s_i)^2$$

$$= \sum_{i=1}^{n}\left[(-2)(s_i - \frac{n+1}{2})\right]^2$$

$$= 4\sum_{i=1}^{n}(s_i - \frac{n+1}{2})^2$$

Dengan menggunakan persamaan (2.3.7), didapatkan :

$$\sum_{i=1}^{n}(d_i+d'_i)^2 = 4\sum_{i=1}^{n}(s_i - \frac{n+1}{2})^2$$

$$= 4\sum_{i=1}^{n}(s_i - \bar{s})^2$$

$$= 4\,\frac{n(n^2-1)}{12}$$

$$= \frac{n(n^2-1)}{3} \qquad \text{....................(3.1.4)}$$

Kemudian akan dicari :

$$\sum_{i=1}^{n}(d_i-d'_i) = \sum_{i=1}^{n}\left[i-s_i-(i-s_{n+1-i})\right]$$

$$= \sum_{i=1}^{n}\left[i-s_i-(n+1-i-s_i)\right]$$

$$= \sum_{i=1}^{n} \left[i-s_i-n-1+i+s_i) \right]$$

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

Sehingga :

$$\sum_{i=1}^{n} (d_i-d'_i)^2 = \sum_{i=1}^{n} \left[2i-(n+1) \right]^2$$

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

$$= 4\sum_{i=1}^{n} (s_i - \bar{s})^2$$

$$= \frac{n(n^2-1)}{3} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.1.5)$$

Kemudian kita perhatikan hubungan berikut :

$$\sum_{i=1}^{n} \left[(d_i+d'_i) + (d_i-d'_i) \right]^2 = 4\sum_{i=1}^{n} d_i^2$$

Maka dengan mensubstitusikan persamaan (3.1.4) dan (3.1.5) kedalam hubungan diatas didapatkan :

$$4\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (d_i+d'_i)^2 + \sum_{i=1}^{n} (d_i-d'_i)^2 + 2\sum_{i=1}^{n} (d_i+d'_i)(d_i-d'_i)$$

$$= \sum_{i=1}^{n} (d_i+d'_i)^2 + \sum_{i=1}^{n} (d_i-d'_i)^2 + 2\sum_{i=1}^{n} d_i^2 - 2\sum_{i=1}^{n} d'_i{}^2$$

$$= \frac{n(n^2-1)}{3} + \frac{n(n^2-1)}{3} + 2\sum_{i=1}^{n} d_i^2 - 2\sum_{i=1}^{n} d'_i{}^2$$

Kemudian :

$$4\sum_{i=1}^{n} d_i^2 - 2\sum_{i=1}^{n} d_i^2 + 2\sum_{i=1}^{n} d'_i{}^2 = \frac{n(n^2-1)}{3} + \frac{n(n^2-1)}{3}$$

$$2\sum_{i=1}^{n} d_i^2 + 2\sum_{i=1}^{n} d'_i{}^2 = 2\,\frac{n(n^2-1)}{3}$$

Sehingga akan didapat :

$$\sum_{i=1}^{n} d_i^2 + \sum_{i=1}^{n} d'_i{}^2 = \frac{n(n^2-1)}{3}$$

$$d + d' = \frac{n(n^2-1)}{3}$$

Jadi didapatkan bahwa untuk setiap himpunan pasangan rank $\{(1,s_1),(2,s_2),\ldots\ldots,(n,s_n)\}$ dengan harga variabel random D sebesar d , maka selalu terdapat himpunan pasangan rank sekawan $\{(1,s_n),(2,s_{n-1}),\ldots\ldots,(n,s_1)\}$ dengan harga variabel random D sebesar d', sehingga :

$$d + d' = \frac{n(n^2-1)}{3}$$

$$\frac{d + d'}{2} = \frac{n(n^2-1)}{6}$$

Dapat disimpulkan bahwa variabel random D simetri terhadap titik $\frac{d + d'}{2} = \frac{n(n^2-1)}{6}$

Karena setiap nilai d dari variabel random D terkait dengan nilai statistik r_s dimana $r_s = 1 - \frac{6\,d}{n(n^2-1)}$, maka dapat disimpulkan :

1. Untuk setiap nilai d, dengan nilai :

$$r_s = 1 - \frac{6\,d}{n(n^2-1)}$$

2. Untuk setiap nilai d', dengan nilai :

$$r'_s = 1 - \frac{6\, d'}{n(n^2-1)}$$

dimana :

$$\frac{d + d'}{2} = \frac{n(n^2-1)}{6} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1.6)$$

Kita uraikan bahwa :

$$r_s = 1 - \frac{6\, d}{n(n^2-1)}$$

$$n(n^2-1)r_s = n(n^2-1) - 6\, d$$

$$6\, d = n(n^2-1) - n(n^2-1)r_s$$

$$d = \frac{n(n^2-1)\ .\ (1-r_s)}{6}$$

Analog diatas maka didapat :

$$d' = \frac{n(n^2-1)\ .\ (1-r'_s)}{6}$$

Sehingga persamaan (3.1.6) dapat ditulis :

$$\frac{\dfrac{n(n^2-1)\ .\ (1-r_s)}{6} + \dfrac{n(n^2-1)\ .\ (1-r'_s)}{6}}{2} = \frac{n(n^2-1)}{6}.$$

$$\frac{n(n^2-1)}{6}\left[(1-r_s) + (1-r'_s)\right] = 2\,\frac{n(n^2-1)}{6}$$

$$(1-r_s) + (1-r'_s) = 2$$

$$2 - (r_s+r'_s) = 2$$

$$1 - \frac{r_s+r'_s}{2} = 1$$

$$\frac{r_s+r'_s}{2} = 0$$

Dapat disimpulkan bahwa untuk setiap r_s , selalu terdapat r'_s sehingga $\frac{r_s+r'_s}{2} = 0$; dengan kata lain statistik r_s simetri terhadap nol.

Contoh (3.1.1)

Distribusi nol eksak dari r_s untuk n = 3.

Pendekatan langsung untuk menentukan nilai U_r, adalah dengan jalan menghitung satu-persatu dari tiap titik sampel $\{(r_1,s_1),(r_2,s_2),(r_3,s_3)\}$. Diasumsikan bahwa sampel bivariat $\{(x_1,y_1),(x_2,y_2),(x_3,y_3)\}$ sudah disusun berdasarkan urutan besar relatif X. Maka didapat :

$r_i = i$

$s_i = s_i$

untuk i = 1,2,3

Dari persamaan (2.3.11) diketahui rumus r_s yaitu :

$$r_s = \frac{12\sum_{i=1}^{n} r_i s_i}{n(n^2-1)} - \frac{3\ (n+1)}{(n-1)}$$

Sehingga untuk n = 3 adalah :

$$r_s = \frac{12\sum_{i=1}^{3} r_i s_i}{3.8} - \frac{3.4}{2}$$

$$= \frac{1}{2}\sum_{i=1}^{3} r_i s_i - 6$$

Tabel untuk n = 3 adalah

titik - sampel	$\sum_{i=1}^{3} i\ s_i$	$r_s = \frac{1}{2}\sum_{i=1}^{3} i\ s_i - 6$
(1,1),(2,2),(3,3)	14	1
(1,1),(2,3),(3,2)	13	0,5
(1,2),(2,1),(3,3)	13	0,5
(1,2),(2,3),(3,1)	11	-0,5
(1,3),(2,1),(3,2)	11	-0,5
(1,3),(2,2),(3,1)	10	-1

Jadi distribusi peluang r_s adalah :

$r_s = r$	-1	-0,5	0,5	+1
U_r	1	2	2	1
$f_{r_s}(r) = \frac{U_r}{3!}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

Berbeda dengan distribusi nol eksak dari r_s yang membutuhkan perhitungan satu-persatu dari tiap-tiap titik sampel, maka distribusi marginal dari masing-masing variabel rank R_i dan S_i, dan distribusi gabungan dari variabel-variabel rank pada sampel tunggalnya (pada pengamatan X saja atau pengamatan Y saja), dapat ditentukan dengan teori kombinatorik.

Contoh untuk sampel tunggal Y yaitu :

himpunan variabel random : $Y_1, Y_2, \ldots\ldots\ldots Y_n$

himpunan variabel rank : $S_1, S_2, \ldots\ldots\ldots S_n$

distribusi marginal :

$$f_{S_i}(s_i) = \frac{1}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3.1.7)$$

distribusi gabungan :

$$f_{S_i,S_j}(s_i, s_j) = \frac{1}{n(n-1)} \quad \ldots\ldots\ldots\ldots\ldots \quad (3.1.8)$$

untuk $s_i, s_j = 1, 2, \ldots\ldots n \quad , \quad s_i \neq s_j$

Maka dari persamaan (2.3.2) dan persamaan (3.1.7), didapatkan :

$$E(S_i) = \sum_{i=1}^{n} s_i \; f(s_i)$$

$$= \sum_{i=1}^{n} s_i \quad \frac{1}{n}$$

$$= \frac{n+1}{2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3.1.9)$$

Dan dari persamaan (2.3.7) dan persamaan (3.1.7), kita dapatkan :

$$\text{var } (S_i) = E\left[\{S_i - E(S_i)\}^2\right]$$

$$= E\left[S_i - \overline{S} \right]^2$$

$$= \sum_{i=1}^{n} (s_i - \overline{s})^2 \ f(s_i)$$

$$= \sum_{i=1}^{n} (s_i - \overline{s})^2 \ \frac{1}{n}$$

$$= \frac{n(n^2-1)}{12} \cdot \frac{1}{n}$$

$$= \frac{n^2-1}{12} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1.10)$$

Sedangkan untuk kovariansi, maka didapatkan bahwa untuk semua $i \neq j$ berlaku :

$$\text{cov}(S_i, S_j) = E\ (S_i, S_j) - E\ (S_i)\ E\ (S_j)$$

$$= \sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} (s_i s_j)\ f(s_i, s_j) - \left[\frac{1}{n} \sum_{i=1}^{n} s_i \right] \left[\frac{1}{n} \sum_{j=1}^{n} s_j \right]$$

$$= \frac{1}{n(n-1)} \sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} i\ j - \frac{1}{n^2} \left[\sum_{i=1}^{n} i \right]^2$$

$$= \frac{1}{n^2(n-1)} \left[n \sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} i\ j - (n-1) \left[\sum_{i=1}^{n} i \right]^2 \right]$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1.11)$$

Akan dicari dahulu nilai $\sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} i\ j$, maka :

$$\sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} i\ j = \sum_{i=1}^{n} \left[\sum_{\substack{j=1 \\ i\neq j}}^{n} i\ j \right]$$

$$= \sum_{i=1}^{n}\left[i \sum_{\substack{j=1 \\ i\neq j}}^{n} j \right]$$

$$= 1\sum_{j=2}^{n} j + 2\sum_{\substack{j=1 \\ j\neq 2}}^{n} j + \ldots\ldots\ldots + n\sum_{\substack{j=1 \\ j\neq n}}^{n} j$$

$$= \left(1\sum_{j=1}^{n} j - 1\right) + \left(2\sum_{j=1}^{n} j - 4\right) + \ldots\ldots\ldots + \left(n\sum_{j=1}^{n} j - n^2\right)$$

$$= \left[1\sum_{j=1}^{n} j + 2\sum_{j=1}^{n} j + \ldots\ldots\ldots + n\sum_{j=1}^{n} j\right] - \sum_{j=1}^{n} j^2$$

$$= \left[1 + 2 + \ldots\ldots\ldots + n\right]\sum_{j=1}^{n} j - \sum_{j=1}^{n} j^2$$

$$= \sum_{i=1}^{n} i \sum_{i=1}^{n} i - \sum_{i=1}^{n} i^2$$

$$= \left[\sum_{i=1}^{n} i\right]^2 - \sum_{i=1}^{n} i^2 \qquad \ldots\ldots\ldots\ldots\ldots(3.1.12)$$

Persamaan (3.1.12) diatas disubstitusikan ke persamaan (3.1.11), maka didapat :

$$\mathrm{cov}(S_i, S_j) = \frac{1}{n^2(n-1)}\left[n\left(\left[\sum_{i=1}^{n} i\right]^2 - \sum_{i=1}^{n} i^2\right) - (n-1)\left[\sum_{i=1}^{n} i\right]^2\right]$$

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

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

$$= -\frac{1}{n^2(n-1)}\left[n\left[\sum_{i=1}^{n} i^2\right] - \left[\sum_{i=1}^{n} i\right]^2\right]$$

$$= -\frac{1}{n^2(n-1)}\left[n\,\frac{n(n+1)(2n+1)}{6} - \frac{n^2(n+1)^2}{4}\right]$$

$$= -\frac{1}{n-1}\left[(n+1)\left[\frac{2n+1}{6} - \frac{n+1}{4}\right]\right]$$

$$= -\frac{n+1}{n-1}\left[\frac{4n+4-3n-3}{12}\right]$$

$$= -\frac{n+1}{n-1}\ \frac{n-1}{12}$$

$$= -\frac{n+1}{12} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1.13)$$

Dengan menggunakan teori kombinatorik seperti diatas, maka untuk variabel rank R_i, akan didapatkan hasil yang sama, yaitu :

$$f_{R_i}(r_i) = \frac{1}{n} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.1.14)$$

$$f_{R_i,R_j}(r_i,r_j) = \frac{1}{n(n-1)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.1.15)$$

$$E(R_i) = \frac{n+1}{2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.1.16)$$

$$Var(R_i) = \frac{(n^2-1)}{12} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.1.17)$$

$$Cov(R_i,R_j) = -\frac{(n+1)}{12} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.1.18)$$

Untuk $r_i, r_j = 1,2,3,\ldots\ldots.n$
$r_i \neq r_j$

Sehingga didapatkan bahwa variabel rank R_i dan S_i mempunyai distribusi yang identik, untuk semua i.

Theorema 3.1.2.

Dibawah hipotesa nol berlaku :

$$E\ (\ r_s \mid H_o\ ) = 0\ , \qquad \text{dan} \quad var\ (\ r_s \mid H_o\ ) = \frac{1}{n-1}$$

Bukti :

Dibawah hipotesa nol, bila pengamatan X dan Y saling bebas maka variabel rank pengamatan R_i dan S_i juga saling bebas untuk semua i dan j. Sehingga diperoleh :

$$E\left[\sum_{i=1}^{n} R_i S_i\right] = \sum_{i=1}^{n} E\ (R_i S_i)$$

$$= \sum_{i=1}^{n} E\ (R_i)\ E\ (S_i)$$

$$= n\ E\ (R_i)\ E\ (S_i)$$

$$= n\ \frac{n+1}{2}\ \frac{n+1}{2}$$

$$= n\ \frac{(n+1)^2}{4} \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3.1.19)$$

Karena $r_s = \dfrac{12 \sum_{i=1}^{n} r_i s_i}{n(n^2-1)} - \dfrac{3\ (n+1)}{(n-1)}$

Maka dengan menggunakan hasil dari persamaan (3.1.19), didapat bahwa ekspektasi r_s dibawah H_o adalah :

$$E\ (\ r_s \mid H_o\ ) = E\left[\frac{12 \sum_{i=1}^{n} R_i S_i}{n(n^2-1)} - \frac{3\ (n+1)}{(n-1)}\right]$$

$$= E\left[\frac{12 \sum_{i=1}^{n} R_i S_i}{n(n^2-1)}\right] - E\left[\frac{3\ (n+1)}{(n-1)}\right]$$

$$= \frac{12}{n(n^2-1)}\ E\left[\sum_{i=1}^{n} R_i S_i\right] - \frac{3\ (n+1)}{(n-1)}$$

$$= \frac{12}{n(n^2-1)}\ \frac{n(n+1)^2}{4} - \frac{3\ (n+1)}{(n-1)}$$

$$= \frac{3n\ (n+1)(n+1)}{n\ (n+1)(n-1)} - \frac{3\ (n+1)}{(n-1)}$$

$$= \frac{3\ (n+1)}{(n-1)} - \frac{3\ (n+1)}{(n-1)}$$

$$= 0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.1.20)$$

Untuk membuktikan var(r_s ¦ H_o) dicari dahulu harga $\mathrm{var}\left[\sum_{i=1}^{n} R_i S_i\right]$ yaitu :

$$\mathrm{var}\left[\sum_{i=1}^{n} R_i S_i\right] = E\left[\left(\sum_{i=1}^{n} R_i S_i\right)^2\right] - \left[E\left(\sum_{i=1}^{n} R_i S_i\right)\right]^2$$

$$= E\left[(R_1S_1+R_2S_2+\ldots.+R_nS_n)^2\right] - \left[E(R_1S_1+R_2S_2+\ldots.+R_nS_n)\right]^2$$

$$= E\left[(R_1S_1)^2+(R_2S_2)^2+\ldots.+(R_nS_n)^2 + (R_1S_1R_2S_2+R_1S_1R_3S_3 + \ldots\ldots+R_1S_1R_nS_n + R_2S_2R_1S_1+R_2S_2R_3S_3+\ldots.+R_2S_2R_nS_n + \ldots\ldots.+R_nS_nR_1S_1 + R_nS_nR_2S_2+\ldots\ldots+R_nS_nR_{n-1}S_{n-1})\right] + \left[E(R_1S_1+R_2S_2+\ldots.+R_nS_n)\right]^2.$$

$$= E\left[(R_1S_1)^2+(R_2S_2)^2+\ldots.+(R_nS_n)^2 + \left[R_1S_1\sum_{j=2}^{n} R_jS_j + R_2S_2\sum_{\substack{j=1\\ j\neq 2}}^{n} R_jS_j +\cdots+ R_nS_n\sum_{\substack{j=1\\ j\neq n}}^{n} R_jS_j\right]\right] - \left[E(R_1S_1+R_2S_2+\ldots.+R_nS_n)\right]^2$$

$$= E\left[(R_1S_1)^2+(R_2S_2)^2+\ldots.+(R_nS_n)^2 + \left[\sum_{i=1}^{n} R_iS_i \sum_{\substack{j=1\\ i\neq j}}^{n} R_jS_j\right]\right] - \left[E(R_1S_1+R_2S_2+\ldots.+R_nS_n)\right]^2$$

$$= E\left[(R_1S_1)^2+(R_2S_2)^2+\ldots.+(R_nS_n)^2 + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\ i<j}}^{n} (R_iS_i)(R_jS_j)\right] - \left[E(R_1S_1+R_2S_2+\ldots.+R_nS_n)\right]^2$$

$$= E\left[(R_1S_1)^2\right]+E\left[(R_2S_2)^2\right]+\ldots+E\left[(R_nS_n)^2\right] +2\sum_{\substack{i=1\\i<j}}^{n}\sum_{j=1}^{n}E\left[(R_iS_i)(R_jS_j)\right] - \left[E(R_1S_1+R_2S_2+\ldots+R_nS_n)\right]^2$$

$$= E\left[(R_1S_1)^2\right]+E\left[(R_2S_2)^2\right]+\ldots+E\left[(R_nS_n)^2\right]+2\sum_{\substack{i=1\\i<j}}^{n}\sum_{j=1}^{n}E\left[(R_iS_i)(R_jS_j)\right] - \left[\left[E(R_1S_1)\right]^2+\left[E(R_2S_2)\right]^2+\ldots+\left[E(R_nS_n)\right]^2 + 2\sum_{\substack{i=1\\i<j}}^{n}\sum_{j=1}^{n}E\left[(R_iS_i)\right]E\left[(R_jS_j)\right]\right]$$

$$= E\left[(R_1S_1)^2\right] + E\left[(R_2S_2)^2\right]+\ldots+E\left[(R_nS_n)^2\right] - \left[\left[E(R_1S_1)\right]^2+\left[E(R_2S_2)\right]^2+\ldots+\left[E(R_nS_n)\right]^2\right]+2\sum_{\substack{i=1\\i<j}}^{n}\sum_{j=1}^{n}E\left[(R_iS_i)(R_jS_j)\right] - 2\sum_{\substack{i=1\\i<j}}^{n}\sum_{j=1}^{n}E\left[(R_iS_i)\right]E\left[(R_jS_j)\right]$$

$$= \left(E\left[(R_1S_1)^2\right]-\left[E(R_1S_1)\right]^2\right) + \left(E\left[(R_2S_2)^2\right]-\left[E(R_2S_2)\right]^2\right) + \ldots + \left(E\left[(R_nS_n)^2\right]-\left[E(R_nS_n)\right]^2\right) + 2\sum_{\substack{i=1\\i<j}}^{n}\sum_{j=1}^{n}\left(E\left[(R_iS_i)(R_jS_j)\right] - E\left[(R_iS_i)\right]E\left[(R_jS_j)\right]\right)$$

$$= \mathrm{var}(R_1S_1)+\ldots+\mathrm{var}(R_nS_n) + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\i<j}}^{n}\mathrm{cov}(R_iS_i,R_jS_j)$$

$$= n\ \mathrm{var}(R_iS_i) + 2\,\frac{n(n-1)}{2}\,\mathrm{cov}(R_iS_i,R_jS_j)$$

$$= n\ \mathrm{var}(R_iS_i) + n(n-1)\ \mathrm{cov}(R_iS_i,R_jS_j) \qquad \ldots\ldots\ldots(3.1.21)$$

Kemudian dicari $\mathrm{var}(R_iS_i)$ dan $\mathrm{cov}(R_iS_i,R_jS_j)$, yaitu :

$$\mathrm{var}(R_iS_i) = E\left[(R_iS_i)^2\right] - \left[E(R_iS_i)\right]^2$$

$$= E\left[(R_i^2)(S_i^2)\right] - \left[E(R_i)E(S_i)\right]^2$$

$$= E\ (R_i^2)E(S_i^2) - \left[\mu_{R_i}\ \mu_{S_i}\right]^2$$

$$= E\ (R_i^2)E(S_i^2) - \left[\mu_{R_i}^2\right]^2$$

$$= E\ (R_i^2)E(S_i^2) - \mu_{R_i}^4 \qquad \ldots\ldots\ldots\ldots(3.1.22)$$

$$\begin{aligned} var(R_i)var(S_i) &= \left(E(R_i^2)-\left[E(R_i)\right]^2\right)\left(E(S_i^2)-\left[E(S_i)\right]^2\right) \\ &= E(R_i^2)E(S_i^2) - E(R_i^2)\left[E(S_i)\right]^2 - \left[E(R_i)\right]^2 \\ &\quad E(S_i^2) + \left[E(R_i)\right]^2\left[E(S_i)\right]^2 \\ &= E(R_i^2)E(S_i^2) - E(R_i^2)\mu_{S_i}^2 - \mu_{R_i}^2E(S_i^2) + \\ &\quad \mu_{R_i}^2\ \mu_{S_i}^2 \\ &= E(R_i^2)E(S_i^2) - \mu_{S_i}^2E(R_i^2) - \mu_{R_i}^2E(S_i^2) + \mu_{R_i}^4 \end{aligned}$$

Jadi didapat :

$$E(R_i^2)E(S_i^2) = var(R_i)var(S_i) + \mu_{S_i}^2E(R_i^2) + \mu_{R_i}^2E(S_i^2) - \mu_{R_i}^4$$

Sehingga persamaan (3.1.22) menjadi :

$$\begin{aligned} var(R_iS_i) &= var(R_i)var(S_i) + \mu_{S_i}^2\ E(R_i^2) + \mu_{R_i}^2\ E(S_i^2) - \\ &\quad \mu_{R_i}^4 - \mu_{R_i}^4 \\ &= var(R_i)var(S_i) + \mu_{S_i}^2\ \left[var(R_i)+\mu_{R_i}^2\right] + \\ &\quad \mu_{R_i}^2\ \left[var(S_i)+\mu_{S_i}^2\right] - 2\ \mu_{R_i}^4 \\ &= var(R_i)var(S_i) + \mu_{S_i}^2\ var(R_i) + \mu_{R_i}^4 + \\ &\quad \mu_{R_i}^2\ var(S_i) + \mu_{R_i}^4 - 2\ \mu_{R_i}^4 \\ &= var(R_i)var(S_i) + \mu_{S_i}^2\ var(R_i) + \mu_{R_i}^2\ var(S_i) \end{aligned}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots(3.1.23)$$

Kemudian akan dicari $cov(R_iS_i, R_jS_j)$ sebagai berikut :

$$cov(R_iS_i, R_jS_j) = E(R_iS_i\ R_jS_j) - E(R_iS_i)\ E(R_jS_j)$$

$$= E(R_iR_j)\ E(\ S_iS_j) - \mu_{R_i}\mu_{S_i} \cdot \mu_{R_i}\mu_{S_i}$$

$$= E(R_iR_j)\ E(\ S_iS_j) - \mu_{R_i}^2\ \mu_{S_i}^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3.1.24)$$

$$\mathrm{cov}(R_i,R_j)\mathrm{cov}(S_i,S_j) = \left[\ E(R_iR_j) - E(R_i)E(R_j)\ \right]\left[E(S_iS_j) - E(S_i)E(S_j)\ \right]$$

$$= \left[E(R_iR_j)-\mu_{R_i}\mu_{R_j}\ \right]\left[E(S_iS_j)-\mu_{S_i}\mu_{S_j}\ \right]$$

$$= E(R_iR_j)E(S_iS_j) - \mu_{S_i}^2\ E(R_iR_j) - \mu_{R_i}^2\ E(S_iS_j) + \mu_{R_i}^2\ \mu_{S_i}^2$$

Jadi didapat :

$$E(R_iR_j)E(S_iS_j) = \mathrm{cov}(R_i,R_j)\mathrm{cov}(S_i,S_j) + \mu_{S_i}^2\ E(R_iR_j) + \mu_{R_i}^2\ E(S_iS_j) - \mu_{R_i}^2\ \mu_{S_i}^2$$

Sehingga persamaan (3.1.24) menjadi :

$$\mathrm{cov}(R_iS_i,R_jS_j) = \mathrm{cov}(R_i,R_j)\mathrm{cov}(S_i,S_j) + \mu_{S_i}^2\ E(R_iR_j) + \mu_{R_i}^2\ E(S_iS_j) - 2\ \mu_{R_i}^2\ \mu_{S_i}^2 \ \ldots\ldots\ldots(3.1.25)$$

Dengan mensubstitusikan persamaan (3.1.23) dan (3.1.25) ke dalam persamaan (3.1.21) , maka didapat :

$$\mathrm{var}\left[\sum_{i=1}^{n} R_iS_i\right]=$$

$$= n\ \left[\mathrm{var}(R_i)\mathrm{var}(S_i) + \mu_{S_i}^2\ \mathrm{var}(R_i) + \mu_{R_i}^2\ \mathrm{var}(S_i)\right] + n(n-1)\left[\mathrm{cov}(R_i,R_j)\mathrm{cov}(S_i,S_j) + \mu_{S_i}^2\ E(R_iR_j) + \mu_{R_i}^2\ E(S_iS_j) - 2\ \mu_{R_i}^2\ \mu_{S_i}^2\ \right]$$

$$= n\ \mathrm{var}(R_i)\mathrm{var}(S_i) + n(n-1)\mathrm{cov}(R_i,R_j)\mathrm{cov}(S_i,S_j) + n\left[\mu_{R_i}^2\ \mathrm{var}(R_i) + \mu_{R_i}^2\ \mathrm{var}(R_i)\right] + n(n-1)\left[\mu_{R_i}^2\ E(R_iR_j) + \mu_{R_i}^2\ E(R_iR_j) - 2\mu_{R_i}^4\right]$$

$$= n\ \mathrm{var}(R_i)\mathrm{var}(S_i) + n(n-1)\mathrm{cov}(R_i,R_j)\mathrm{cov}(S_i,S_j) + n\left[2\ \mu_{R_i}^{2}\ \mathrm{var}(R_i)\right] + n(n-1)\left[2\ \mu_{R_i}^{2}\ E(R_iR_j) - 2\ \mu_{R_i}^{4}\right]$$

$$= n\ \mathrm{var}(R_i)\mathrm{var}(S_i) + n(n-1)\mathrm{cov}(R_i,R_j)\mathrm{cov}(S_i,S_j) + 2n\mu_{R_i}^{2}\left[\mathrm{var}(R_i)+(n-1)\left[E(R_iR_j)-\mu_{R_i}^{2}\right]\right]$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.1.26)$$

Akan dicari nilai $\mathrm{var}(R_i) + (n-1)\left[E(R_iR_j)-\mu_{R_i}^{2}\right]$, yaitu dengan menggunakan persamaan (3.1.16) dan (3.1.17), maka:

$$\mathrm{var}(R_i)+(n-1)\left[E(R_iR_j)-\mu_{R_i}^{2}\right] = \frac{n^2-1}{12} + (n-1)\left[E(R_iR_j) - \frac{(n+1)^2}{4}\right]$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.1.27)$$

Akan dicari dahulu $E(R_iR_j)$ yaitu :

$$E(R_iR_j) = \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} r_i r_j\ f(r_i r_j)$$

$$= \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} i\ j\ \frac{1}{n(n-1)}$$

Dengan menggunakan persamaan (3.1.12), didapatkan :

$$E(R_iR_j) = \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} i\ j\ \frac{1}{n(n-1)}$$

$$= \left(\left[\sum_{i=1}^{n} i\right]^2 - \sum_{i=1}^{n} i^2\right)\frac{1}{n(n-1)}$$

$$= \left(\left[\frac{n(n+1)}{2}\right]^2 - \frac{n(n+1)(2n+1)}{6}\right)\frac{1}{n(n-1)}$$

$$= \frac{1}{n(n-1)}\left[\frac{3n^2(n+1)^2}{12} - \frac{2n(n+1)(2n+1)}{12}\right]$$

$$= \frac{1}{12\ n(n-1)}\left[3n^2(n+1)^2 - 2n(n+1)(2n+1)\right]$$

$$= \frac{3n^3+2n^2-3n-2}{12(n-1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1.28)$$

Sehingga persamaan (3.1.27) akan menjadi :

$$\mathrm{var}(R_i)+(n-1)\left[E(R_iR_j)-\mu_{R_i}^2\right] =$$

$$= \frac{n^2-1}{12} + (n-1)\left[\frac{3n^3+2n^2-3n-2}{12(n-1)} - \frac{(n+1)^2}{4}\right]$$

$$= \frac{n^2-1}{12} + \left[\frac{(n-1)\ \ 3n^2+5n+2}{12} - \frac{3(n-1)(n+1)^2}{12}\right]$$

$$= \frac{n^2-1}{12} + \frac{(n-1)}{12}\left[3n^2+5n+2-3n^2-6n-3\right]$$

$$= \frac{n^2-1}{12} + \frac{(n-1)}{12}\left[-(n+1)\right]$$

$$= \frac{n^2-1}{12} - \frac{n^2-1}{12}$$

$$= 0$$

Karena $\mathrm{var}(R_i)+(n-1)\left[E(R_iR_j)-\mu_{R_i}^2\right] = 0$, maka persamaan (3.1.26) adalah :

$$\mathrm{var}\left[\sum_{i=1}^{n} R_iS_i\right] = n\ \mathrm{var}(R_i)\mathrm{var}(S_i) + n(n-1)\mathrm{cov}(R_i,R_j)$$

$$\mathrm{cov}(S_i,S_j) + 2\ n\ \mu_{R_i}^2\ [\ 0\ ]$$

$$= n\mathrm{var}(R_i)\mathrm{var}(S_i)+n(n-1)\mathrm{cov}(R_i,R_j)\mathrm{cov}(S_i,S_j)$$

Dengan menggunakan persamaan (3.1.17) dan persamaan (3.1.18) maka didapat :

$$\mathrm{var}\left[\sum_{i=1}^{n} R_iS_i\right] = n\ \frac{(n^2-1)}{12}\ \frac{(n^2-1)}{12} + n(n-1)\left[\frac{-(n+1)}{12}\right]\left[\frac{-(n+1)}{12}\right]$$

$$= \frac{n(n+1)(n-1)(n+1)(n-1)}{144} + \frac{n(n+1)(n-1)(n+1)}{144}$$

$$= \frac{n(n-1)(n+1)^2\ \ (n-1+1)}{144}$$

$$= \frac{n^2(n-1)(n+1)^2}{144} \qquad \ldots\ldots\ldots\ldots\ldots(3.1.29)$$

Kemudian, untuk mencari variansi r_s dibawah hipotesa nol dengan cara sebagai berikut :

$$\mathrm{var}\ (\ r_s|H_o\ ) = E(r_s^{\,2}) - \left[E(r_s)\right]^2$$

Dari persamaan (3.1.20) diketahui $E(r_S) = 0$, sehingga:

$$\mathrm{var}\ (\ r_s|H_o\ ) = E(r_s^{\,2}) \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3.1.30)$$

Akan dicari $E(r_s^{\,2})$ dengan cara melihat kembali persamaan (2.3.10), yaitu :

$$r_s = \frac{12 \displaystyle\sum_{i=1}^{n} r_i s_i - \frac{n(n+1)^2}{4}}{n(n^2-1)}$$

$$r_s^{\,2} = \left[\frac{12}{n(n^2-1)}\right]^2 \left[\sum_{i=1}^{n} r_i s_i - \frac{n(n+1)^2}{4}\right]^2$$

$$= \left[\frac{12}{n(n^2-1)}\right]^2 \left[\left[\sum_{i=1}^{n} r_i s_i\right]^2 - \frac{2n(n+1)^2}{4}\sum_{i=1}^{n} r_i s_i + \frac{n^2(n+1)^4}{16}\right]$$

$$= \left[\frac{12}{n(n^2-1)}\right]^2 \left[\left[\sum_{i=1}^{n} r_i s_i\right]^2 - \frac{n(n+1)^2}{2}\sum_{i=1}^{n} r_i s_i + \frac{n^2(n+1)^4}{16}\right]$$

Sehingga didapat :

$$E(r_s^{\,2}) = E\left\{\left[\frac{12}{n(n^2-1)}\right]^2 \left(\left[\sum_{i=1}^{n} r_i s_i\right]^2 - \frac{n(n+1)^2}{2}\sum_{i=1}^{n} r_i s_i + \frac{n^2(n+1)^4}{16}\right)\right\}$$

$$E(r_s^{\,2}) = \left[\frac{12}{n(n^2-1)}\right]^2 \left\{E\left[\sum_{i=1}^{n} r_i s_i\right]^2 - \frac{n(n+1)^2}{2}\, E\left[\sum_{i=1}^{n} r_i s_i\right]\right.$$

$$+ \frac{n^2(n+1)^4}{16}\Bigg\}$$

Dari persamaan 3.1.19 diketahui bahwa :

$$E\left[\sum_{i=1}^{n} R_i S_i\right] = \frac{n(n+1)^2}{4}$$

Maka :

$$E(r_s^2) = \left[\frac{12}{n(n^2-1)}\right]^2 \left\{ E\left[\sum_{i=1}^{n} r_i s_i\right]^2 - \frac{n(n+1)^2}{2}\,\frac{n(n+1)^2}{4} + \frac{n^2(n+1)^4}{16}\right\}$$

$$= \left[\frac{12}{n(n^2-1)}\right]^2 \left\{ E\left[\sum_{i=1}^{n} r_i s_i\right]^2 - \frac{n^2(n+1)^4}{16}\right\}$$

$$= \left[\frac{12}{n(n^2-1)}\right]^2 \left\{ E\left[\sum_{i=1}^{n} r_i s_i\right]^2 - \left(E\left[\sum_{i=1}^{n} r_i s_i\right]\right)^2\right\}$$

$$= \left[\frac{12}{n(n^2-1)}\right]^2 \text{var} \left[\sum_{i=1}^{n} r_i s_i\right] \quad \ldots\ldots\ldots\ldots\ldots(3.1.31)$$

Sehingga persamaan (3.1.30) dapat dituliskan :

$$\text{var } (r_s | H_o) = E(r_s^2)$$

$$= \left[\frac{12}{n(n^2-1)}\right]^2 \text{var} \left[\sum_{i=1}^{n} r_i s_i\right] \quad \ldots\ldots\ldots(3.1.32)$$

Dengan mensubstitusikan persamaan (3.1.29) ke persamaaan (3.1.32) , maka :

$$\text{var } (r_s | H_o) = \frac{144}{n^2(n-1)^2(n+1)^2}\,\frac{n^2(n-1)(n+1)^2}{144}$$

$$= \frac{1}{(n-1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.1.33)$$

Dari keseluruhan perhitungan diatas, maka dapat disimpulkan bahwa dibawah hipotesis nol berlaku :

$$E\ (r_s | H_o) = 0$$

$$\text{var}\ (r_s | H_o) = \frac{1}{(n-1)}$$

3.2 Distribusi Nol Asimptotik Dari r_s

Dari persamaan (2.3.11) diketahui bahwa :

$$r_s = \frac{12 \sum_{i=1}^{n} r_i s_i}{n(n^2-1)} - \frac{3\ (n+1)}{(n-1)}$$

Sampel random bivariat $\{(x_1,y_1),(x_2,y_2),\ldots\ldots\ldots,(x_n,y_n)\}$ tersebut sudah disusun berdasarkan ururan besar relatif komponen X dari terkecil hingga terbesar.

Jadi didapat bahwa :

$r_i = i$

$s_i = S_i$

untuk $i = 1,2,3,\ldots\ldots\ldots,n$

Dari persamaan (2.3.11), distribusi sampling dari r_s hanya bergantung kepada variabel random $\sum_{i=1}^{n} i\ S_i$. Sedangkan ekspektasi dan variansi dari $\sum_{i=1}^{n} i\ S_i$ ini, dibawah hipotesa nol,sudah diturunkan dalam persamaan (3.1.19) dan (3.1.29) yaitu :

$$E\left[\sum_{i=1}^{n} R_i S_i\right] = n\ \frac{(n+1)^2}{4} \quad \ldots\ldots\ldots\ldots\ldots\ldots(3.2.1)$$

$$\text{var}\left[\sum_{i=1}^{n} R_i S_i\right] = \frac{n^2(n-1)(n+1)^2}{144} \quad \ldots\ldots\ldots\ldots\ldots (3.2.2)$$

Atau dapat dituliskan

$$\sum_{i=1}^{n} i\ S_i \sim N\left[n\ \frac{(n+1)^2}{4}\ ,\ \frac{n^2(n-1)(n+1)^2}{144}\right]$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.2.3)$$

Maka dibawah hipotesa nol, statistik :

$$r_s = \frac{12\sum_{i=1}^{n} r_i s_i}{n(n^2-1)} - \frac{3\ (n+1)}{(n-1)}$$

, mempunyai distribusi asimptotik normal dengan :

$$E\ (r_s) = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.2.4)$$

$$\text{var}\ (r_s) = \frac{1}{(n-1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.2.5)$$

Atau dapat dituliskan :

$$r_s \sim N\left[0\ ,\ \frac{1}{(n-1)}\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.3.6)$$

Sehingga variabel random normal baku yang biasa digunakan untuk uji kebebasan adalah :

$$Z = \sqrt{(n-1)}\ r_s \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.3.7)$$

Pendekatan ini cukup baik walaupun n hanya sebesar 10.

3.3 Uji Signifikansi Dari r_s

Untuk sampel kecil, harga-harga kritis r_s ditabelkan pada tabel A pada lampiran untuk n dari 4 hingga 30. Jika suatu harga observasi r_s sama dengan atau melampaui harga yang ditabelkan, harga observasi tersebut signifikan pada tingkat yang ditunjukkan. Untuk sampel besar digunakan persamaan (3.2.7) untuk n lebih dari 30 dan harga-harga kritis terdapat pada tabel B pada lampiran. H_o dapat ditolak jika harga suatu harga Z sama dengan atau melampaui harga yang ditabelkan, pada tingkat yang ditunjukkan.

Hipotesis-hipotesis :

1. Uji dua sisi.

   H_o : $\mu_1 = \mu_2$ (X dan Y saling bebas)

   H_1 : $\mu_1 \neq \mu_2$ (Ada pertalian antara X dan Y)

2. Uji satu sisi.

   H_o : $\mu_1 \leq \mu_2$ (X dan Y saling bebas)

   H_1 : $\mu_1 > \mu_2$ (Ada pertalian langsung antara X dan Y)

3. Uji satu sisi.

   H_o : $\mu_1 \geq \mu_2$ (X dan Y saling bebas)

   H_1 : $\mu_1 < \mu_2$ (Ada pertalian invers antara X dan Y)

Hal lain yang perlu dicatat adalah koefisien korelasi rank Spearman mengukur derajat hubungan diantara rank dari variabel X dan Y, dan bukan derajat hubungan diantara variabel itu sendiri.