MODUL 7

Inferensi Nonparametrik

Prof. Dr. Zanzawi Soejoeti

PENDAHULUAN

Statistika nonparametrik adalah sekumpulan prosedur inferensi yang berlaku di bawah bentuk distribusi yang sangat luas. Istilah *inferensi nonparametrik* diturunkan dari kenyataan bahwa manfaat prosedur-prosedur ini tidak memerlukan model distribusi tertentu (misalnya normal) untuk populasinya. Dalam uji hipotesis, statistik penguji nonparametrik biasanya hanya memanfaatkan sifat-sifat sederhana data sampelnya, seperti tanda aljabar hasil pengukuran, urutannya atau frekuensi kategorinya.

Sifat umum ini tidak memerlukan adanya skala numerik pengukurannya, memperlebar atau mempersempit skala itu tidak mengubahnya. Akibatnya, distribusi nol suatu statistik penguji nonparametrik dapat ditentukan tanpa melihat bentuk distribusi populasi yang melandasinya. Oleh karena alasan ini, prosedur-prosedur itu juga dinamakan uji distribusi bebas. Sifat distribusi bebas ini merupakan sifat yang sangat menguntungkan karena itu banyak statistik yang lebih senang mengatakan prosedur-prosedur itu sebagai prosedur distribusi bebas daripada prosedur nonparametrik.

Apabila data merupakan hasil pengukuran pada suatu skala numerik dan anggapan bahwa data itu merupakan sampel random dari suatu populasi normal berlaku maka tentu saja prosedur parametrik akan lebih efisien dalam arti bahwa umumnya prosedur uji akan mempunyai kekuatan yang lebih besar dan interval kepercayaan akan lebih pendek dari saingannya yang nonparametrik. Kemauan untuk menganggap lebih banyak tentang bentuk populasinya akan menyebabkan prosedur inferensi yang lebih baik, tetapi risikonya jika anggapan itu tidak sesuai kita akan menggunakan prosedur yang tidak benar dan memperoleh kesimpulan yang salah. Jadi, pemilihan antara prosedur parametrik dan nonparametrik harus didasarkan atas pertimbangan antara kurangnya efisiensi dan tingkat perlindungan terhadap tidak dipenuhinya anggapan-anggapan itu.

KEGIATAN BELAJAR 1

Inferensi dengan Satu Sampel dan Dua Sampel Independen

A. INFERENSI DENGAN SATU SAMPEL

Definisi median:

> Misalkan, X suatu variabel random dengan distribusi peluang kontinu. Median X adalah bilangan M sehingga
> $$P(X \leq M) = P(X \geq M) = 0{,}5$$

Distribusi suatu variabel random kontinu X dikatakan *simetrik* terhadap suatu nilai b jika $P(X \leq b-c) = P(X \geq b-c)$ untuk setiap $c \geq 0$. Jadi, suatu variabel random X yang simetrik terhadap suatu nilai b akan mempunyai median M dan mean μ keduanya sama dengan b.

Untuk suatu distribusi yang menceng, kerap kali median dipandang lebih informatif daripada mean karena mean sangat dipengaruhi oleh beberapa nilai yang ekstrem.

Sebagaimana kita selalu menaksir mean suatu populasi dengan mean sampel yang diambil dari populasi itu maka kita juga akan menaksir median suatu populasi dengan median sampel yang diambil dari populasi itu.

Misalkan, $\{x_1, ..., x_n\}$ sampel random berukuran n dari suatu populasi. Didefinisikan

$x_{(1)}$ = observasi terkecil dalam sampel

$x_{(2)}$ = observasi kedua terkecil dalam sampel

…

$x_{(i)}$ = observasi ke i dalam sampel

…

$x_{(n)}$ = observasi terbesar dalam sampel

Nilai-nilai observasi (pengamatan) $x_{(1)} \leq x_{(2)} \leq ..., \leq x_{(n)}$ dinamakan statistik berurut dari sampel $\{x_1, ..., x_n\}$.

Definisi:

> Median x̃ suatu sampel $\{x_1,...,x_n\}$ didefinisikan sebagai bilangan
>
> $$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}}, & \text{jika n ganjil} \\ \frac{1}{2}\left[x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right], & \text{jika } n \text{ genap} \end{cases}$$

1. **Interval Kepercayaan (Sampel Kecil)**

Untuk menghitung interval kepercayaan (1-α)100% bagi median M suatu populasi yang berdistribusi kontinu, dipilih statistik berurut $X_{(r)}$ dan $X_{(s)}$ sehingga $P\left(X_{(r)} \le M \le X_{(s)}\right) = 1-\alpha$.

Metode memilih r dan s sebagai fungsi n dan α adalah sebagai berikut.

Bilangan bulat r dipilih dari tabel binomial kumulatif sedemikian sehingga:

$$P(Y < r) = P(Y \le r-1) = \sum_{i=0}^{r-1} \binom{n}{i} \left(\tfrac{1}{2}\right)^n = \frac{\alpha}{2}$$

maka

$$P\left(X_{(r)} \le M \le X_{(n-r+1)}\right) = 1-\alpha.$$

Bilangan bulat r yang memenuhi peluang di atas tidak dapat ditemukan untuk α yang ditentukan. Dalam praktik, kita pilih r yang menghasilkan α yang dekat dengan nilai yang kita inginkan. Contoh berikut akan menggambarkan prosedur ini.

Contoh 7.1

Dipunyai 10 observasi upah mingguan tukang kayu di PQR sebagai suatu sampel random sebagai berikut (dalam rupiah).

$x_1 = 9.850$ $x_4 = 20.118$ $x_7 = 8.641$ $x_{10} = 11.200$

$x_2 = 7.223$ $x_5 = 10.670$ $x_8 = 22.087$

$x_3 = 6.421$ $x_6 = 4.274$ $x_9 = 8.400$

Dari data ini diperoleh median sampel

$$\tilde{x} = \frac{1}{2}\left[x_{(5)} + x_{(6)}\right] = \frac{1}{2}(8.641 + 9.850) = 9.245,5$$

Akan kita cari interval kepercayaan 90% bagi median M. Dari Tabel I (tabel peluang binomial kumulatif) untuk n = 10 dan p = 0,5: $P(Y \leq 2) = 0,055$ dan $P(Y \leq 3) = 0,172$. Jadi, dipilih r = 2 maka diperoleh $(1-\alpha) = 1 - 2(0,055) = 0,89$, yang akan menghasilkan interval kepercayaan 89% bagi M sebagai berikut.

$$X_{(2)} = 6.421 \leq M \leq 20.118 = X_{(9)}$$

Prosedur di atas dapat juga diaplikasikan untuk distribusi yang diskrit, hanya saja peluangnya ditulis sebagai:

$$P\left(X_{(r)} \leq M \leq X_{(n-r+1)}\right) \geq 1 - \alpha$$

2. Interval Kepercayaan (Sampel Besar, $n \geq 20$)

Jika Y berdistribusi binomial dengan n trial dan peluang sukses p = 0,5 maka dengan pendekatan normal diperoleh:

Mean (Y) = n p = 0,5 n

dan

Deviasi standar

$$Y = \sqrt{np(1-p)} = \sqrt{(0,5)^2 n}$$

sehingga dengan faktor koreksi = 0,5 diperoleh:

$$\begin{aligned} P(Y \leq r-1) &= P(Y \leq r - 1 + 0,5) = P(Y \leq r - 0,5) \\ &= P\left[\frac{Y - 0,5n}{\sqrt{(0,5)^2 n}} \leq \frac{r - 0,5 - 0,5n}{\sqrt{(0,5)^2 n}}\right] \\ &= P\left[Z \leq Z^*\right] \end{aligned}$$

dengan $Z^* = \dfrac{r - (n+1)0,5}{0,5\sqrt{n}}$ dan $Z \sim N(0;1)$.

Jika kita pilih $Z^* = -Z_{\alpha/2}$ maka $P(Y \leq r-1) \approx P\left(Z \leq Z^*\right) = \alpha/2$

untuk $r = (n+1)0,5 - Z_{\alpha/2}.(0,5)\sqrt{n}$

Tetapi di sini biasanya r tidak bulat. Maka kita pilih r sebagai bilangan bulat terbesar yang lebih kecil atau sama dengan (≤) $(n+1)0,5 - Z_{\alpha/2}.(0,5)\sqrt{n}$ sehingga diperoleh interval kepercayaan pendekatan $(1-\alpha)\,100\%$ sebagai $P\left(X_{(r)} \leq M \leq X_{(n-r+1)}\right) \approx 1-\alpha$

Contoh 7.2

Dari suatu sampel dengan n = 36 diperoleh observasi-observasi sebagai berikut.

x_1	=	120	x_{10}	=	140	x_{19}	=	160	x_{28}	=	114
x_2	=	210	x_{11}	=	275	x_{20}	=	215	x_{29}	=	335
x_3	=	195	x_{12}	=	130	x_{21}	=	410	x_{30}	=	120
x_4	=	85	x_{13}	=	90	x_{22}	=	125	x_{31}	=	160
x_5	=	173	x_{14}	=	175	x_{23}	=	150	x_{32}	=	70
x_6	=	155	x_{15}	=	260	x_{24}	=	159	x_{33}	=	104
x_7	=	116	x_{16}	=	180	x_{25}	=	100	x_{34}	=	195
x_8	=	225	X_{17}	=	325	X_{26}	=	215	x_{35}	=	140
x_9	=	168	x_{18}	=	80	x_{27}	=	170	x_{36}	=	80

Jika data di atas kita urutkan maka akan diperoleh data berurutan sebagai berikut.

70	104	130	160	180	225
80	114	140	160	195	260
80	116	140	168	195	275
85	120	150	170	210	325
90	120	155	173	215	335
100	125	159	175	215	410

Misalkan, kita ingin menaksir median M dengan interval kepercayaan 95%. Di sini $\alpha = 0,05$ dan $Z_{\alpha/2} = 1,96$ sehingga

$(n+1)0,5 - Z_{\alpha/2}.(0,5)\sqrt{n} = 37(0,50) - 1,96(0,5)(6) = 12,62$ dan r = 12 adalah bilangan bulat $\leq 12,62$. Maka interval kepercayaan (kira-kira) 95% adalah $X_{(12)} \leq M \leq X_{(25)}$

Dari data berurut di atas dapat kita amati bahwa $X_{(12)}$ dan $X_{(25)} = 180$. Sehingga interval kepercayaan itu menjadi $125 \leq M \leq 180$

B. INFERENSI DENGAN DUA SAMPEL INDEPENDEN

Masalah membandingkan dua populasi berdasarkan sampel-sampel random independen telah pernah kita pelajari dalam kuliah Metode Statistika I di bawah anggapan distribusi populasinya normal dan bervariansi sama. Prosedur parametrik yang kita pelajari di sana berdasarkan statistik t. Di sini kita akan mempelajari prosedur nonparametrik untuk masalah yang sama. Prosedur ini pertama kali diusulkan oleh Wilcoxon, dan versi alternatif yang ekuivalen dengan itu diusulkan secara independen oleh Mann dan Whitney.

Untuk studi pembandingan dua perlakuan A dan B, sekumpulan $n = n_A + n_B$ unit eksperimen dibagi secara random menjadi dua kelompok masing-masing berukuran n_A dan n_B. Perlakuan A digunakan untuk n_A unit, dan perlakuan B digunakan untuk n_B unit yang lain. Pengukuran hasilnya kita tulis sebagai

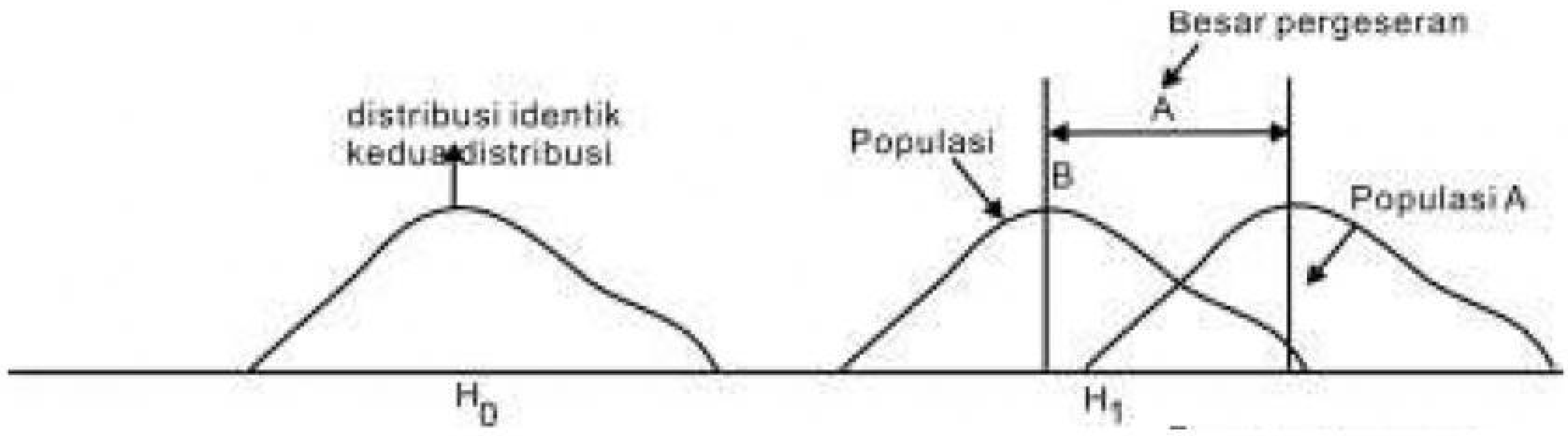
Perlakuan A: $x_{11}, x_{12}, \ldots, x_{1n_A}$

Perlakuan B: $x_{21}, x_{22}, \ldots, x_{2n_B}$

Dua perlakuan ini merupakan sampel-sampel random independen dari dua populasi. Dengan menganggap bahwa hasil yang lebih besar menunjukkan perlakuan yang lebih baik. Kita ingin menguji hipotesis nol bahwa tidak ada perbedaan antara dua efek perlakuan versus alternatif satu sisi bahwa perlakuan A lebih efektif daripada perlakuan B. Dalam penataan nonparametrik ini dan distribusinya dianggap kontinu.

Model	:	Kedua distribusi kontinu
Hipotesis	:	
H_0	:	kedua distribusi populasi identik

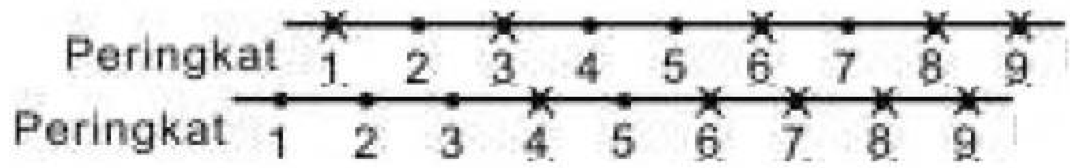
H_1	:	Distribusi populasi A tergeser ke kanan dari distribusi populasi B.



Gambar 7.1
Representasi H_0 dan H_1 dalam Besar Pergeseran Δ

Perhatikan bahwa tidak ada anggapan tentang *bentuk* distribusi populasinya. Keadaan ini sangat berbeda dengan prosedur parametrik, di mana kita menganggap bahwa distribusi populasinya normal dengan variansi sama. Gambar 7.1 melukiskan hipotesis H_0 dan H_1 di atas.

Konsep dasar yang melandasi uji jumlah peringkat sekarang dapat dijelaskan dengan alur penalaran intuitif sebagai berikut. Misalkan, 2 himpunan observasi itu digambarkan pada diagram yang sama, menggunakan tanda yang berbeda (titik dan silang) untuk menunjukkan sumbernya yang berbeda. Di bawah H_0, sampel-sampel itu berasal dari populasi yang sama sehingga 2 himpunan titik-titik itu bercampur dengan baik, misalnya observasi-observasi yang lebih besar, lebih banyak yang berasal dari sampel pertama, kita dapat menyimpulkan bahwa mungkin populasi A tergeser ke kanan dari populasi B. Dua keadaan ini dilukiskan dalam Gambar 7.2, di mana dalam setiap kasus himpunan titik-titik gabungan diberi nomor berurut dari kiri ke kanan. Nomor-nomor ini dinamakan *peringkat sampel gabungan*. Dalam Gambar 7.2a baik peringkat besar maupun kecil bercampur bagi elemen-elemen kedua sampel itu, sedangkan dalam Gambar 7.2b kebanyakan peringkat yang lebih besar bagi elemen-elemen sampel pertama. Maka, dengan menggunakan jumlah peringkat yang berkaitan dengan sampel pertama sebagai statistik penguji, ini berarti, nilai yang besar untuk statistik ini harus mencerminkan keadaan populasi pertama terletak ke kanan dari populasi kedua.



Gambar 7.2
Gabungan Dua Sampel dan Peringkat Sampel Gabungan

Guna menentukan daerah penolakan dengan tingkat signifikansi tertentu, kita harus memperhatikan distribusi jumlah-peringkat di bawah hipotesis nol. Konsep ini ditelusuri dalam Contoh 7.3 di mana ukuran-ukuran sampel kecil dipelajari untuk menghindari deretan pencacahan yang panjang.

Contoh 7.3

Guna menentukan apakah suatu biji hibrida baru menghasilkan tanaman bunga daripada varietas yang saat ini popular, seseorang ahli hortikultura menanam 2 biji hibrida baru dan 3 biji yang saat ini popular dalam suatu petak kebun. Setelah tanaman tumbuh dewasa, dicatat

Hibrida baru A	31,8	39,1	
Varietas sekarang B	35,5	27,6	21,3

Apakah data ini menunjukkan dengan kuat bahwa hibrida baru menghasilkan tanaman bunga?

Kita ingin menguji hipotesis nol

H_0 : Populasi A dan B identik

Versus hipotesis alternatif

H_1 : Populasi A bergeser dari populasi B ke arah nilai-nilai yang lebih besar

Dalam uji jumlah peringkat, kedua sampel itu dicampur dan diletakkan bersama dan selanjutnya diberi peringkat dari kecil ke besar.

Observasi Berurut Sampel Gabungan	**21,3**	**27,6**	**31,8**	**35,5**	**39,1**
Peringkat	1	2	3	4	5
Varietas	B	B	A	B	A

Jumlah peringkat untuk A: $W_A = 3 + 5 = 8$

Jumlah peringkat untuk B: $W_B = 1 + 2 + 4 = 7$

Oleh karena pengukuran (observasi) yang lebih tinggi untuk varietas A cenderung mendukung H_1, daerah penolakan uji kita harus terdiri dari nilai-nilai W_A yang besar:

H_0 ditolak jika $W_A \geq C$

Untuk menentukan nilai kritis C sedemikian sehingga peluang kesalahan tipe I terkontrol pada tingkat α yang ditentukan, kita perhatikan distribusi peluang W_A di bawah H_0. Jika dua sampel berasal dari populasi yang sama, setiap pasang bilangan bulat yang diambil dari $\{1,2,3,4,5\}$ berpeluang sama akan menjadi peringkat bagi dua hasil pengukuran A. Ada $\binom{5}{2} = 10$ pasang yang dapat kita peroleh sehingga tiap himpunan peringkat yang mungkin mempunyai peluang $\frac{1}{10} = 0{,}1$ di bawah H_0. Himpunan-himpunan peluang ini disajikan dalam Tabel 7.1 dengan nilai-nilai W_A yang bersesuaian.

Tabel 7.1
Himpunan Peringkat untuk Perlakuan A dengan Ukuran Sampel $n_A = 2$ dan $n_B = 3$.

Peringkat A	**Jumlah Peringkat W_A**	**Peluang**
1,2	3	0,1
1,3	4	0,1
1,4	5	0,1
1,5	6	0,1
2,3	5	0,1
2,4	6	0,1
2,5	7	0,1
3,4	7	0,1
3,5	8	0,1
4,5	9	0,1
	Jumlah	1,0

Distribusi nol W_A segera dapat diperoleh dari Tabel 7.1 dengan mengumpulkan nilai-nilai yang identik (Tabel 7.2).

Tabel 7.2
Distribusi Jumlah Peringkat W_A untuk Ukuran Sampel $n_A = 2$ dan $n_B = 3$.

Nilai W_A	3	4	5	6	7	8	9
Peluang	0,1	0,1	0,2	0,2	0,2	0,1	0,1

Nilai pengamatan $W_A = 8$ mempunyai peluang signifikansi $P_{H_0}(W_A \geq 8) = 0,1 + 0,1 = 0,2$ dengan perkataan lain, kita harus berani mempunyai peluang kesalahan tipe I sebesar 0,2 untuk nilai pengamatan W_A akan menghasilkan kesimpulan menolak H_0. Uji jumlah peringkat membawa kita ke kesimpulan bahwa fakta tidak cukup kuat untuk menolak H_0. Perhatikan sekiranya pengukuran A menerima peringkat-peringkat tertinggi dari 4 dan 5, masih akan diperlukan tingkat signifikansi $\alpha = 0,1$ untuk menolak H_0.

Dengan mengacu Contoh 7.1, kita rumuskan prosedur uji jumlah peringkat secara umum.

Uji jumlah peringkat Wilcoxon
Misalkan, $X_{11}, X_{12}, ..., X_{1n_A}$ dan $X_{21}, X_{22}, ..., X_{1n_B}$
Sampel-sampel random independen masing-masing dari populasi kontinu A dan B. Guna menguji H_0 : populasi-populasi itu identik:

a. Berilah peringkat sampel gabungan $n = n_A + n_B$ observasi dalam urutan naik.
b. Hitunglah jumlah peringkat sampel pertama W_A
c. (i) Untuk H_1: populasi A tergeser ke kanan dari populasi B; pasanglah daerah penolakannya pada ekor atas W_A.
   (ii) Untuk H_1: populasi A tergeser ke kanan dari populasi B; pasanglah daerah penolakannya pada ekor bawah W_A.
   (iii) Untuk H_1: populasi-populasi itu berbeda; pasanglah daerah penolakannya pada kedua ekor W_A dan mempunyai peluang sama.

Menentukan distribusi nol statistik jumlah peringkat secara langsung menjadi sangat membosankan jika ukuran sampel bertambah besar. Oleh karena itu, telah disediakan tabel untuk distribusi nol statistik ini untuk sampel-sampel kecil, dan distribusi pendekatannya tersedia untuk sampel-

sampel besar. Untuk menjelaskan penggunaan tabel statistik jumlah peringkat Wilcoxon, pertama-tama kita catat jumlah peringkat W_A dan W_B .

Jumlah dua jumlah peringkat $W_A + W_B$ adalah tetap, yang merupakan jumlah bilangan-bilangan bulat 1, 2, …, n, dengan n adalah ukuran sampel gabungan.

Misalnya, dalam Tabel 7.1.

$$\begin{aligned} W_A + W_B &= (3+5)+(1+2+4) \\ &= 1+2+3+4+5 \\ &= 15 \end{aligned}$$

Dengan demikian, suatu uji yang menolak H_0 untuk nilai-nilai W_A yang besar adalah ekuivalen dengan uji yang menolak H_0 untuk nilai-nilai W_B yang kecil. Kita dengan mudah dapat menggunakan W_B sebagai statistik penguji dan menata daerah penolakan, pada ekor bawah. Oleh karena itu, kita selalu dapat memusatkan perhatian pada jumlah peringkat sampel yang lebih kecil dan memasang daerah penolakan pada ekor bawah atau atas bergantung apakah hipotesis alternatif menyatakan distribusi populasi yang bersangkutan bergeser ke kiri atau ke kanan.

Kedua, distribusi masing-masing statistik jumlah peringkat W_A dan W_B adalah simetrik. Yakni, W_A simetrik terhadap $n_A\left(n_A + n_B + 1\right)/2$ dan W_B simetrik terhadap $n_B\left(n_A + n_B + 1\right)/2$. Tabel 7.2 melukiskan simetrik distribusi W_A itu untuk $n_A = 2,\ n_B = 3$. Mahasiswa dapat menentukan titik simetrinya. Sifat-sifat statistik ini sangat membantu dalam membentuk tabel distribusinya.

1. Tabel Statistik Jumlah Peringkat Wilcoxon

Statistik uji jumlah peringkat Wilcoxon diambil sebagai W_S = jumlah peringkat sampel yang kecil dalam peringkat sampel gabungan. Jika ukuran sampel sama, ambillah jumlah peringkat sampel yang mana saja. Tabel statistik jumlah peringkat Wilcoxon memberikan baik peluang ekor atas atau ekor bawah, yaitu:

$P\left(W_S \geq X\right)$ = peluang ekor atas

$P\left(W_S \leq X^*\right)$ = peluang ekor bawah

Dari sifat simetri distribusi itu, peluang di atas akan sama jika X dan X^* berjarak sama dari pusat. Tabel itu memuat nilai-nilai X^* yang berkaitan dengan nilai-nilai X pada ekor atas. Misalnya, dengan ukuran sampel 3 dan 7, tabel menunjukkan $P(W_S \geq 25) = 0,033 = P(W_S \leq 8)$. Langkah-langkah yang harus diikuti jika menggunakan tabel jumlah peringkat Wilcoxon dalam melakukan uji jumlah peringkat adalah:

Gunakan jumlah peringkat sampel yang kecil W_S sebagai statistik penguji. Jika ukuran sampel sama ambil salah satu jumlah peringkat sebagai W_S. Jika H_1 menyatakan bahwa populasi yang berkaitan dengan W_S :

a. tergeser ke kanan dari populasi yang lain maka gunakan daerah penolakan yang berbentuk $W_s \geq C$ dan ambil C sebagai nilai X yang terkecil sehingga $P \leq \alpha$.
b. tergeser ke kiri maka gunakan daerah penolakan yang berbentuk $W_s \leq C$ dan ambil C sebagai nilai X^* yang terbesar sehingga $P \leq \alpha$.
c. tergeser ke salah satu arah maka gunakan daerah penolakan yang berbentuk $[W_s \leq C_1 \text{ atau } W_s \geq C_2]$ dan membaca C_1 dari kolom X^* dan C_2 dari kolom X sehingga $P \leq \dfrac{\alpha}{2}$.

Contoh 7.4

Dua formasi geologi akan kita bandingkan dalam hal kekayaan kandungan mineralnya. Kandungan mineral 7 contoh biji yang diambil dari formasi 1 dan 5 contoh biji yang diambil dari formasi 2 diukur dengan analisis kimia. Diperoleh data sebagai berikut:

	Kandungan Mineral						
Formasi 1	7,6	11,1	6,8	9,8	4,9	6,1	15,1
Formasi 2	4,7	6,4	4,1	3,7	3,9		

Apakah data memberikan dukungan yang kuat bahwa formasi 1 mempunyai kandungan mineral yang lebih tinggi dari formasi 2? Ujilah dengan α mendekati 0,05.

Untuk menggunakan uji jumlah peringkat, pertama-tama kedua sampel itu kita gabung dan kita beri peringkat. Selanjutnya, kita tentukan jumlah peringkat untuk sampel kedua, yakni sampel yang ukurannya lebih kecil. Pengamatan dalam sampel kedua dan peringkatnya digarisbawahi supaya dapat diidentifikasi dengan cepat.

Sampel Gabungan (Berurut)	<u>3,7</u>	<u>3,9</u>	<u>4,1</u>	<u>4,7</u>	4,9	6,1	<u>6,4</u>	6,8	7,6	9,8	11,1	15,1
Peringkat	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	5	6	<u>7</u>	8	9	10	11	12

Nilai pengamatan statistik jumlah peringkat adalah
$W_S = 1 + 2 + 3 + 4 + 7 = 17$

Kita ingin menguji hipotesis nol bahwa distribusi kedua populasi itu identik versus hipotesis alternatif bahwa populasi kedua, bersesuaian dengan nilai W_S, terletak di sebelah kiri dari yang pertama. Dengan demikian, daerah penolakannya ada pada ekor bawah dari W_S.

Dari tabel jumlah peringkat Wilcoxon, dengan sampel yang kecil berukuran 5 dan yang besar berukuran 7, diperoleh $P(W_S \leq 21) = 0,037$ dan $P(W_S \leq 22) = 0,053$. Jadi, dengan $\alpha = 0,053$ kita mempunyai daerah penolakan $W_S \leq 22$. Oleh karena nilai hitungan $W_S = 17$ jatuh di daerah ini, hipotesis nol ditolak dengan $\alpha = 0,053$. Bahkan hipotesis nol itu akan ditolak dengan α sekecil $P(W_S \leq 17) = 0,005$.

Contoh 7.5

Bahan tahan api diuji dengan menyalakan selembar kertas pada lipatan pakaian yang dipakai seseorang. Kerusakannya diukur (dalam cm). Diperoleh data (untuk dua jenis bahan) sebagai berikut.

Bahan A	5,7	7,3	7,6	6,0	6,5
Bahan B	4,9	7,4	5,3	4,6	6,2

Apakah data menunjukkan gejala yang kuat adanya perbedaan tahan api kedua bahan tersebut. Ujilah dengan α sekitar 0,05.

Ukuran kedua sampel itu sama sehingga kita dapat mengambil jumlah peringkat sampel yang mana pun sebagai statistik penguji. Kita hitung jumlah peringkat untuk sampel kedua sebagai berikut.

Nilai Berurut	<u>4,6</u>	<u>4,9</u>	<u>5,3</u>	5,7	6,0	<u>6,2</u>	6,5	7,3	<u>7,4</u>	7,6
Peringkat	<u>1</u>	<u>2</u>	<u>3</u>	4	5	<u>6</u>	7	8	<u>9</u>	10

$$W_S = 1 + 2 + 3 + 6 + 9 = 21$$

Oleh karena hipotesis alternatif dua sisi, daerah penolakannya meliputi kedua ekor dari W_S. Dari tabel kita peroleh

$$P(W_S \geq 37) = 0,028 = P(W_S \leq 18)$$

Jadi, dengan $\alpha = 0,056$, daerah penolakannya adalah $\{W_S \geq 37\}$ atau $\{W_S \leq 18\}$. Kita lihat bahwa nilai hitungan kita $\{W_S = 21\}$ tidak jatuh dalam daerah penolakan sehingga hipotesis nol tidak ditolak pada $\alpha = 0,056$.

2. Pendekatan Sampel Besar

Jika ukuran sampel yang digunakan besar, distribusi statistik jumlah peringkat di bawah hipotesis nol mendekati distribusi normal. Dengan demikian, uji hipotesis dilakukan dengan menggunakan tabel distribusi normal. Misalkan, ukuran sampel n_A dan n_B keduanya besar, dan W_A menunjukkan jumlah peringkat untuk sampel berukuran n_A maka W_A dapat dianggap mendekati distribusi normal.

Di bawah H_0:

$$\text{Mean } W_A = \frac{n_A(n_A + n_B + 1)}{2}$$

$$\text{Variansi } W_A = \frac{n_A n_B(n_A + n_B + 1)}{12}$$

Statistik jumlah peringkat

$$Z = \frac{W_A - n_A\left(n_A + n_B + 1\right)/2}{\sqrt{n_A n_B \left(n_A + n_B + 1\right)/12}}$$

Mendekati distribusi $N\left(0;1\right)$ jika H_0 benar dan sampel besar.

Daerah penolakan untuk statistik Z dapat ditentukan dengan menggunakan tabel distribusi normal standar.

Untuk melukiskan besar kesalahan yang terkandung dalam pendekatan ini, kita pandang kasus $n_A = 9$ dan $n_B = 10$. Daerah penolakan satu sisi dengan pendekatan normal adalah

$$\frac{W_A - 9\left(20\right)/2}{\sqrt{9 \times 10 \times 20/12}} = \frac{W_A - 90}{12,247} \geq 1,64$$

yang dapat disederhanakan menjadi $W_A \geq 110,1$. Dari tabel statistik, jumlah peringkat Wilcoxon, yaitu diperoleh $P(W_A \geq 110) = 0,056$ dan $P(W_A \geq 111) = 0,047$ yang sangat dekat dengan $\alpha = 0,05$. Tentu saja, kesalahan pendekatan akan menjadi lebih kecil jika ukuran sampel bertambah besar.

3. Bagaimana Menangani Observasi yang Sama

Dalam contoh-contoh yang lalu, semua observasi dalam sampel gabungan semuanya berbeda. Dengan demikian, peringkatnya dapat ditentukan dengan mudah karena kurang tepatnya skala pengukuran, sering kali nilai-nilai pengukuran terulang dalam satu atau kedua sampelnya. Sebagai contoh, marilah kita perhatikan himpunan data berikut.

Sampel 1	20	24	22	24	26
Sampel 2	26	28	26	30	18

Kita peroleh sampel gabungan berurut sebagai berikut.

$$18 \quad 20 \quad 22 \quad \underbrace{24 \quad 24} \quad \underbrace{26 \quad 26 \quad 26} \quad 28 \quad 30$$

Di sini terdapat dua observasi berulang (sama), yakni yang pertama mempunyai dua unsur, dan yang kedua tiga unsur. Dua posisi yang ditempati 24 seharusnya berperingkat 4 dan 5 maka masing-masing observasi ini kita beri peringkat rata-ratanya, yakni $\frac{4+5}{2}=4{,}5$. Demikian juga, untuk tiga observasi 26, yang seharusnya mendapat peringkat 6, 7, dan 8, masing-masing kita beri peringkat $\frac{6+7+8}{3}=7$. Setelah memberi peringkat rata-rata kepada observasi yang sama dan peringkat biasa kepada observasi yang lain, selanjutnya statistik jumlah peringkat kita hitung seperti biasanya, tetapi jika terdapat observasi yang sama dalam sampel-sampel yang kecil, tabel jumlah peringkat Wilcoxon tidak lagi tepat dan kita perlukan sedikit perubahan. Namun, hal ini tidak kita bicarakan di sini. Kita hanya akan membicarakan perubahan untuk sampel besar.

Untuk sampel-sampel besar, berlaku pendekatan normal untuk statistik jumlah peringkat. Rumus mean sampel seperti di atas, tetapi rumus variansinya berbeda, yakni:

$$\text{variansi } W_A = \frac{n_A n_B (n+1)}{12}\left[1-\frac{\sum_{j=1}^{\ell} a_j\left(a_j^2-1\right)}{n\left(n^2-1\right)}\right]$$

dengan $n = n_A + n_B$

ℓ = banyak observasi yang sama

a_j = banyak elemen sama yang ke j; j = 1, 2, …, ℓ

4. Interval Kepercayaan Berdasarkan Uji Peringkat

Penting untuk ditekankan di sini bahwa prosedur nonparametrik tidak hanya berlaku untuk uji hipotesis, tetapi dapat juga digunakan untuk menghitung interval kepercayaan. Hubungan antara interval kepercayaan dengan uji berbagai hipotesis nol, yang pernah kita pelajari, dapat digunakan untuk menentukan interval kepercayaan yang berkaitan dengan uji nonparametrik yang kita pelajari dalam modul ini. Uji dua ekor akan memberikan interval kepercayaan yang kita bicarakan berikut ini.

Dari Gambar 7.1 di depan, kita dapat melihat bahwa hipotesis alternatif menyatakan bahwa distribusi populasi A bergeser ke kanan dari populasi B sebesar Δ. Kita perhatikan suatu sifat yang menyatakan bahwa jika Δ dikurangkan dari setiap observasi populasi A maka observasi terubah $X_{1i} - \Delta$ akan mempunyai distribusi yang sama, seperti observasi populasi B. Hal ini menunjukkan bahwa kita dapat mengurangkan suatu nilai Δ yang mungkin, kemudian kita gunakan uji jumlah peringkat Wilcoxon dua sisi. Suatu interval kepercayaan dapat diperoleh, yakni terdiri dari semua nilai-nilai Δ ini dengan hipotesis nol tidak ada perbedaan perlakuan tidak ditolak. Untuk ini kita lakukan langkah-langkah sebagai berikut.

a. Untuk semua $n_A - n_B$ pasang observasi yang mungkin, satu dari sampel A dan satu dari sampel B, kita tentukan selisih $x_{1i} - x_{2j}$ yang merupakan nilai taksiran untuk Δ.
b. Selisih-selisih ini kita urutkan dari kecil ke besar.
c. Interval kepercayaan untuk besar pergeseran Δ adalah interval terkecil ke d sampai terbesar ke d termasuk titik-titik ujungnya. Nilai-nilai d dan koefisien kepercayaannya dapat dilihat dari tabel yang tersedia.

Jika ada observasi yang sama, interval itu paling tidak mempunyai tingkat kepercayaan seperti yang dinyatakan.

Contoh 7.6

Hitunglah interval kepercayaan untuk besar pergeseran bagi data dalam Contoh 7.4. Hanya beberapa selisih terkecil dan terbesar yang harus dihitung, misalnya $X_{17} - X_{24} = 15{,}1 - 3{,}9 = 11{,}2$. Dengan cara seperti itu kita peroleh:

Terkecil	-1,5	-0,3	0,2	0,4	0,8	1,0	1,2	1,2
Terbesar	-	7,2	7,4	8,7	10,4	11,0	11,2	11,4

Dari tabel yang tersedia, untuk ukuran sampel yang lebih kecil 5 dan yang lebih besar 7, kita peroleh d = 6 untuk $1 - \alpha = 0{,}952$. Jadi, [1, 0 ; 7, 4] adalah interval kepercayaan 95,2% untuk Δ. Karena interval ini tidak memuat nol, dapat kita simpulkan bahwa dua formasi itu mempunyai muatan mineral yang berbeda secara signifikan.

Jika diagram titik mengungkapkan bahwa dua distribusi tampak berbeda dalam sesuatu hal selain pergeserannya, tentu saja interval kepercayaan untuk besar pergeseran tidak sesuai di sini, misalnya satu sampel menunjukkan lebih memencar dari yang lain, kecuali berbeda dalam letaknya.

LATIHAN

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

1) Sampel-sampel random yang independen masing-masing berukuran $n_A = 4$ dan $n_B = 2$ diambil dari dua populasi yang kontinu.

   Hitunglah banyak himpunan peringkat yang mungkin yang berkaitan dengan *sampel yang kecil dalam peringkat sampel digabung. Berikan nilai peluang kepada himpunan peringkat ini* di bawah hipotesis nol bahwa kedua populasi itu identik

   Carilah distribusi statistic $W_s =$ jumlah peringkat sampel yang kecil.

   Periksalah bahwa peluang-peluang ekornya sesuai dengan nilai-nilai yang telah ditabelkan
2) Gunakan tabel statistik jumlah peringkat Wilcoxon yang tersedia guna menentukan
   a. $P(W_s \geq 42)$ jika $n_1 = 5; n_2 = 7$.
   b. $P(W_s \leq 25)$ jika $n_1 = 6; n_2 = 6$.
   c. $P(W_s \geq 81 \text{ atau } W_s \leq 45)$ jika $n_1 = 10; n_2 = 7$.
   d. Titik C sehingga $P(W_s \leq C)$=0,036 jika $n_1 = 8; n_2 = 4$.
   e. Titik C_1 dan C_2 sehingga $P(W_s \geq C_2) = P(W_s \leq C_1) 0,05$ jika $n_1 = 3; n_2 = 9$
3) Dalam suatu studi tentang kapasitas kognitif binatang menyusui, 19 ekor kera yang berumur sama dibagi secara random menjadi dua kelompok, masing-masing dengan 10 dan 9 ekor. Kelompok-kelompok itu dilatih dengan 2 metode yang berbeda untuk mengingat suatu perangsang akustik. Skor yang diperoleh kera-kera itu pada tes berikutnya adalah sebagai berikut.

Metode I	167	149	137	178	179	155	164	104	151	150
Metode II	98	127	140	103	116	105	100	95	131	

Apakah data dengan kuat menunjukkan adanya perbedaan dalam kemampuan mengingat kera-kera yang dilatih dengan dua metode itu? Gunakan uji jumlah peringkat Wilcoxon dengan α dekat dengan 10%. Hitunglah interval kepercayaan hampir 90% untuk Δ.

4) Uji median dua sampel. Seperti uji jumlah peringkat Wilcoxon, uji median adalah prosedur nonparametrik untuk uji kesamaan dua distribusi populasi berdasarkan sampel-sampel random independen. Masukan sample-sampel random berukuran n_A dan n_B diambil dari populasi kontinu A dan B. Dengan mengurutkan semua $n = n_A + n_B$ observasi, tentukan median X^* sampel digabung. Selanjutnya, kita cacah banyak observasi sampel pertama yang terletak di bawah median X^*, dan gunakan T untuk menunjukkan cacah ini. Di bawah hipotesis nol, kita mengharapkan akan memperoleh kira-kira setengah dari observasi sampel pertama dan setengah dari sampel kedua ada pada salah satu sisi median X^*. Pergeseran populasi A ke sebelah kanan B ditunjukkan dengan nilai-nilai T yang kecil.

   Jika ukuran-ukuran sampel itu besar, statistik penguji T mendekati distribusi normal di bawah H_0, dengan

   $$\text{mean} = \frac{n_A}{2} \text{ ; variansi} = \frac{n_A\ n_B}{4(n_A + n_B)}$$

   Fakta ini dapat dimanfaatkan guna melakukan uji berdasarkan sampel-sampel besar dengan alternatif satu sisi atau dua sisi.

   Kita dapat juga membuat Tabel 2 × 2 sebagai berikut (untuk mudahnya, kita hanya memandang kasus jika n genap).

	Banyak Observasi		**Jumlah**
	Di Bawah X^*	**Di Atas X^***	
Sampel Pertama	T	$n_A - T$	n_A
Sampel Kedua	$\frac{n}{2} - T$	$\frac{n}{2} - n_A + T$	n_B
Jumlah	$\frac{n}{2}$	$\frac{n}{2}$	n

Jelas tampak struktur serupa dengan tabel kategori* 2×2. Maka, untuk sampel-sampel besar dan alternatif 2 sisi, uji median dapat dilakukan sebagai uji X^2 dalam tabel kategori* 2×2.

Untuk penerapan uji median, misalkan uji ketangkasan mengemudi mobil diberikan kepada kelompok mahasiswa putri dan kelompok mahasiswa putra. Semua subjek percobaan itu adalah mahasiswa suatu universitas yang kira-kira berumur sama dan mempunyai pengalaman mengemudi selama tiga tahun (minimal). Diperoleh skor seperti berikut ini (skor tinggi menunjukkan kualitas mengemudi yang baik).

Mahasiswa:	**Skor Uji Ketangkasan Mengemudi**													
Putri	28	53	39	27	41	68	27	28	45	48	65	78		
Putra	32	35	61	43	82	44	78	38	85	63	46	30	47	57

Dengan menggunakan uji median, apakah perbedaan keterampilan mengemudi antara mahasiswa putri dan mahasiswa putra signifikan? Hitung interval kepercayaan 95% untuk Δ.

RANGKUMAN

Istilah infersi nonparametrik diturunkan dari kenyataan bahwa manfaat prosedur-prosedur ini tidak memerlukan model distribusi tertentu untuk populasinya.

Inferensi nonparametrik bermanfaat khususnya jika sifat yang kita pelajari tidak dapat dicatat dalam skala numerik, tetapi urutannya dapat kita amati.

Kita pelajari interval kepercayaan untuk median berdasarkan satu sampel, baik sampel kecil maupun sampel besar (pendekatan normal).

Selanjutnya kita pelajari inferensi dengan dua sampel independen. Ini meliputi jumlah peringkat Wilcoxon (sampel kecil dan sampel besar) dam interval kepercayaan untuk besar pergeseran Δ.

TES FORMATIF 1

Pilihlah satu jawaban yang paling tepat!

I. Suatu sampel 10 roket menunjukkan waktu tingkat pertama pembakaran (dalam menit) sebagai berikut.

6,2 8,7 7,5 14,1 6,3 7,1 8,1 7,3 6,7 7,0

1) Median waktu pembakaran data ini adalah ....
   A. 7,1
   B. 7,2
   C. 7,3
   D. 7,4

2) Interval kepercayaan (1 - α) 100% untuk median waktu pembakaran (jika α kira-kira 2%) adalah ....
   A. (6,2; 14,1)
   B. (5,6; 13,7)
   C. (4,6; 12,5)
   D. (4,2; 12,2)

3) Dalam soal nomor 2 koefisien kepercayaan yang eksak adalah ....
   A. 0,727
   B. 0,791
   C. 0,812
   D. 0,978

II. Dua merek minuman ringan dibandingkan kadar kalori per liter. Data berikut merupakan hasil uji suatu sampel random dengan 10 liter tiap-tiap merek. Berdasarkan data ini, gunakan tingkat signifikansi 10 % untuk menguji apakah ada perbedaan antara kedua merek itu.

	Kadar Kalori (per liter)									
Merek 1	14	15	11	12	12	16	14	11	13	12
Merek 2	11	12	12	10	14	10	12	12	12	10

4) Rumuskan hipotesis nol dan alternatifnya.
   A. $H_0 : \Delta = 0$
      $H_1 : \Delta = 0$
   B. $H_0 : \Delta = 0$
      $H_1 : \Delta \neq 0$
   C. $H_0 : \Delta > 0$
      $H_1 : \Delta = 0$
   D. $H_0 : \Delta < 0$
      $H_1 : \Delta = 0$

5) Tuliskan daerah kritis uji ini
   A. $W_s \geq 128$ atau $W_s \leq 82$
   B. $W_s \leq 129$ atau $W_s \geq 85$
   C. $W_s \leq 135$ atau $W_s \geq 75$
   D. $W_s \geq 136$ atau $W_s \leq 74$

6) Statistik penguji jumlah peringkat Wilcoxon adalah ....
   A. $105 = 92$
   B. $105 = 97$
   C. $W_s = 116{,}5$
   D. $W_s = 129{,}5$

7) Kesimpulannya adalah ....
   A. H_0 diterima pada tingkat signifikasi 10%
   B. H_1 ditolak pada tingkat signifikasi 5%
   C. H_0 ditolak pada tingkat signifikasi 10%
   D. H_0 diterima pada tingkat signifikasi 5%

III. Seorang ahli penata lalu-lintas ingin melihat apakah penataan lampu stop sepanjang jalan protokol akan mengubah lama waktu perjalanan. Dia mengambil sampel random dengan 20 kendaraan sebelum perubahan dan sampel random dengan 25 kendaraan sesudah perubahan. Diperoleh data sebagai berikut (dalam menit).

Sebelum Perubahan						Sesudah Perubahan					
7,8	5,2	6,3	4,1	8,2	5,7	6,4	5,9	4,8	3,8	5,2	7,2
6,8	7,4	8,3	6,4	4,9	6,0	6,3	5,7	6,7	4,3	3,9	5,8
5,4	8,1	5,5	7,2	5,7	6,9	7,0	4,1	8,6	5,2	8,1	4,2
9,4	6,8					5,0	7,7	4,7	3,6	6,0	4,5
						5,5					

8) Dengan pendekatan normal jumlah peringkat sampel 1 (sebelum perubahan) W_1 mendekati normal dengan mean sama dengan ....
   A. 440
   B. 460
   C. 480
   D. 500

9) Seperti dalam soal nomor 8 maka variansi W_1 sama dengan ....
   A. 191,7
   B. 212,6
   C. 234,7
   D. 241,6

10) Nilai statistik Z, jika H_0 benar adalah ....
   A. 540,5
   B. 550,5
   C. 560,5
   D. 570,5

Cocokkanlah jawaban Anda dengan Kunci Jawaban Tes Formatif 1 yang terdapat di bagian akhir modul ini. Hitunglah jawaban yang benar. Kemudian, gunakan rumus berikut untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 1.

$$\text{Tingkat penguasaan} = \frac{\text{Jumlah Jawaban yang Benar}}{\text{Jumlah Soal}} \times 100\%$$

Arti tingkat penguasaan: 90 - 100% = baik sekali
80 - 89% = baik
70 - 79% = cukup
< 70% = kurang

Apabila mencapai tingkat penguasaan 80% atau lebih, Anda dapat meneruskan dengan Kegiatan Belajar 2. **Bagus!** Jika masih di bawah 80%, Anda harus mengulangi materi Kegiatan Belajar 1, terutama bagian yang belum dikuasai.

KEGIATAN BELAJAR 2

Inferensi dengan Dua Sampel Berpasangan

Inferensi dengan dua sampel berpasangan sudah pernah kita lakukan (lihat Mata Kuliah Metode Statistik I) menggunakan distribusi t berdasarkan selisih observasi berpasangan itu. Untuk sampel-sampel kecil inferensi ini menganggap bahwa selisih observasi berpasangan merupakan sampel random dari suatu populasi normal.

Di sini akan kita pelajari dua uji nonparametrik, *uji tanda* dan *uji peringkat bertanda Wilcoxon*, yang dapat digunakan dengan aman bagi selisih berpasangan jika anggapan distribusi normal diragukan. Selanjutnya akan kita bicarakan juga sedikit tentang interval kepercayaan. Pokok bahasan ini akan kita tutup dengan pembicaraan tentang koefisien korelasi peringkat. Struktur data hasil percobaan pembandingan berpasangan ditunjukkan dalam Tabel 7.3 berikut dengan observasi pasangan $ke-i$ ditulis $\left(X_{1i}, X_{2i}\right)$.

Tabel 7.3
Struktur Data Berpasangan

Pasangan	1	2	...	n
Perlakuan A	X_{11}	X_{12}	...	X_{1n}
Perlakuan B	X_{21}	X_{22}	...	X_{2n}
Selisih D=(A-B)	D_1	D_2	...	D_n

A. INFERENSI DENGAN DUA SAMPEL BERPASANGAN

1. Uji tanda

Uji nonparametrik yang intuitif paling mudah adalah uji tanda. Sebagaimana ditunjukkan oleh namanya, uji tanda didasarkan atas tanda aljabar selisih D_i. Statistik pengujinya adalah:

S = Banyak pasangan yang menunjukkan perlakuan A menghasilkan data lebih tinggi daripada perlakuan B;
= Banyak tanda positif di antara selisih $D_i, i = 1, 2, ..., n$

Jika dua pengaruh (efek) perlakuan benar-benar sama maka selisih $D_i, i = 1, 2, ..., n$ berkemungkinan sama akan positif atau negatif. Demikian juga jika pengukuran dilakukan dengan skala kontinu, tidak akan terjadi hasil pengukuran yang identik dalam suatu pasangan. Dengan demikian, dapat dirumuskan hipotesis nol sebagai berikut:
$H_0 : P(+) = 0,5 = P(-)$.

Dengan mengatakan tanda positif sebagai sukses maka statistik penguji S adalah banyak sukses dalam n trial. Oleh karena itu, di bawah H_0 statistik S berdistribusi binomial (n; 0,5). Jika hipotesis alternatif mengatakan bahwa perlakuan A mempunyai pengaruh lebih besar dari perlakuan B, yang dapat kita artikan $P(+) > 0,5$ maka nilai-nilai S yang besar harus ada dalam daerah penolakan. Jika alternatif itu dua sisi, yakni $H_1 : P(+) \neq 0,5$ maka harus digunakan uji dua sisi.

Contoh 7.7

Dilakukan uji tahan pakai untuk membandingkan busi mobil model baru dengan model lama. Suatu sampel dengan 12 mobil dari model sedan kecil sampai model stasion yang cukup besar dilibatkan dalam studi ini. Jarak tempuh km per liter bensin dicatat untuk tiap mobil dengan menggunakan busi model lama dan busi model baru. Hasil pengukurannya dituangkan dalam Tabel 7.4. berikut ini.

Tabel 7.4
Jarak Tempuh Km Per Liter Bensin untuk 12 Pasang Mobil

Mobil	1	2	3	4	5	6	7	8	9	10	11	12
Busi Baru (A)	26,4	10,3	15,8	16,5	32,5	8,3	22,1	30,1	12,9	12,6	27,3	9,4
Busi Baru (B)	24,3	9,8	16,9	17,2	30,5	7,9	22,4	28,6	13,1	11,6	25,5	8,6
Selisih (A-B)	+2,1	+0,5	-1,1	-0,7	+2,0	+0,4	-0,3	+1,5	-0,2	+1,0	+1,8	+0,8

Memperhatikan selisih (A-B), kita lihat ada delapan tanda tambah dalam sampel berukuran n = 12 itu. Jadi nilai pengamatan statistik uji tanda adalah $S = 8$. Dalam pengujian
H_0: tidak ada perbedaan antara A dan B atau $P(+) = 0,5$
H_1 : A lebih baik daripada B atau $P(+) > 0,5$

kita akan menolak H_0 untuk nilai-nilai S yang besar. Dari tabel distribusi binomial dengan n = 12 dan p = 0,5 kita peroleh $P(S \geq 9) = 0{,}073$ dan $P(S \geq 10) = 0{,}019$. Jika kita ingin mengendalikan α di bawah 0,05, harus digunakan daerah penolakan $S \geq 10$. Nilai S yang diamati, $S = 8$, tidak terletak dalam daerah penolakan maka dengan tingkat signifikansi 0,019 data tidak mendukung kenaikan jarak tempuh. Peluang signifikansi dari nilai pengamatan itu adalah $P(S \geq 8) = 0{,}194$.

Sebenarnya, penerapan uji tanda tidak memerlukan nilai numerik selisih itu. Berapa banyak tanda aljabar yang ada dapat diperoleh hanya dengan memandang datanya. Di samping sedikitnya hitungan numerik yang diperlukan uji tanda mempunyai penerapan yang cukup luas. Meskipun jika sekiranya sifat yang akan kita ukur tidak dapat diukur, tetapi dengan sesuatu cara masih dapat kita katakan mana yang lebih besar sehingga kita dapat mengatakan (A – B) positif atau negatif. Hanya informasi inilah yang kita perlukan untuk melakukan uji tanda.

Untuk sampel-sampel besar, uji tanda dapat dilakukan dengan menggunakan pendekatan distribusi normal untuk distribusi binomial. Dengan n besar, distribusi binomial b(n; 0,5) mendekati distribusi normal dengan mean $n/2$ dan deviasi standar $\sqrt{n/4}$.

Pendekatan sampel besar statistik uji tanda di bawah H_0 :

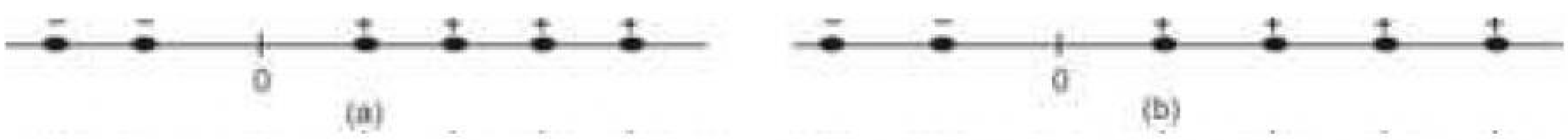
$$Z = \frac{S - n/2}{\sqrt{n/4}}$$

Mendekati distribusi N(0; 1).

Jika dua pengukuran yang kita buat dalam suatu pasangan sama maka selisihnya tidak positif dan tidak negatif. Karena itu, pengukuran yang sama seperti ini kita keluarkan dari sampel, dan mengakibatkan ukuran sampelnya berkurang. Misalnya jika suatu sampel dengan n = 20 pasang menghasilkan 10 tanda positif dan 6 tanda negatif, sedangkan yang 4 nol (pengukuran sama) maka uji tanda dilakukan dengan ukuran sampel $n^* = 20 - 4 = 16$ dan $S = 10$.

2. Uji Peringkat Bertanda Wilcoxon

Telah kita sebutkan di atas bahwa uji tanda dapat menjangkau data ordinal yang sifatnya dalam suatu pasangan dapat dibandingkan tanpa mengukurnya dengan skala numerik. Tetapi jika terdapat pengukuran numerik, uji tanda akan kehilangan informasi cukup banyak karena yang dimanfaatkan hanya tanda aljabar selisih, bukan besarnya selisih itu. Coba kita lihat dan bandingkan dua himpunan selisih berpasangan yang dituangkan dalam diagram titik dalam Gambar 7.3 berikut ini.



Gambar 7.3
Dua Gambar Selisih Berpasangan dengan Banyak Tanda + yang Sama tetapi Dengan Lokasi Berbeda bagi Distribusinya.

Dalam kedua kasus itu terdapat n = 6 titik data dengan empat bertanda aljabar positif sehingga uji tanda akan menghasilkan kesimpulan yang identik, tetapi titik-titik dalam Gambar 7.3 (b) menunjukkan pergeseran ke arah sisi positif yang lebih banyak karena selisih-selisih yang positif terletak ke kanan lebih jauh dari nol dibandingkan selisih-selisih negatif yang tidak terlalu jauh ke kiri dari nol. Seharusnya kita tidak memberi bobot yang sama pada semua tanda positif, seperti yang kita lakukan dalam uji tanda, tetapi kita harus memberi bobot yang lebih besar pada tanda positif yang terletak lebih jauh dari nol. Konsep inilah yang melandasi uji peringkat bertanda.

Dalam uji peringkat bertanda, selisih-selisih pasangan itu diurutkan menurut nilai numeriknya tanpa memandang tanda aljabarnya (nilai mutlaknya), selanjutnya peringkat yang bersesuaian dengan selisih yang positif dijumlahkan untuk digunakan sebagai statistik pengujinya. Guna melukiskan hal ini, kita lihat kembali data jarak kilometer yang diberikan dalam Contoh 7.7. Nilai-nilai selisih tertuang dalam baris terakhir Tabel 7.7. Kita berikan peringkat pada selisih-selisih ini setelah nilai mutlaknya kita urutkan dari kecil ke besar. Selanjutnya kita hitung statistik peringkat bertanda T^+, kita peroleh

$$\begin{aligned} T^+ &= \text{jumlah peringkat yang berkaitan dengan selisih positif} \\ &= 3+4+6+7+9+10+11+12 \\ &= 62 \end{aligned}$$

Selisih	2,1	0,5	-1,1	-0,7	2,0	0,4	-0,3	1,5	-0,2	1,0	1,8	0,8
Nilai mutlak berurut	0,2	0,3	0,4	0,5	0,7	0,8	1,0	1,1	1,5	1,8	2,0	2,1
Peringkat	1	2	3	4	5	6	7	8	9	10	11	12
Tanda	-	-	+	+	-	+	+	-	+	+	+	+

Jika hipotesis nol bahwa tidak ada perbedaan pengaruh perlakuan itu benar maka nilai selisih $\mathrm{D}_1, \mathrm{D}_2, ..., \mathrm{D}_n$ dapat dipandang sebagai suatu sampel random dari suatu populasi yang simetrik terhadap nol. Di samping itu, hipotesis alternatif bahwa "perlakuan A lebih baik" menyatakan bahwa distribusi itu bergeser dari nol ke arah nilai-nilai positif. Di bawah H_1, tidak hanya akan dijumpai lebih banyak tanda aljabar positif, tetapi juga tanda-tanda positif akan cenderung berkaitan dengan peringkat-peringkat yang lebih besar. Dengan demikian, diharapkan T^+ akan besar di bawah alternatif satu sisi, dan kita pilih daerah penolakan pada ekor atas nilai-nilai T^+.

> Langkah-langkah dalam uji peringkat bertanda
>
> (a) Hitung selisih $\mathrm{D}_i = X_{1i} - X_{2i}, 1, 2, ..., n$
>
> (b) Berikan peringkat dengan menyusun nilai-nilai mutlak D_i dalam urutan naik. Catat juga tanda-tanda yang bersesuaian dengan nilai itu.
>
> (c) Hitung statistik peringkat bertanda $T^+ =$ jumlah peringkat selisih D_i yang positif.
>
> (d) Tentukan daerah penolakan pada ekor atas atau bawah atau kedua ekor statistik T^+, sesuai dengan pernyataan apakah pengaruh perlakuan A lebih tinggi, rendah atau berbeda dengan pengaruh perlakuan B dalam hipotesis alternatif.

Beberapa nilai peluang ekor T^+ di bawah H_0 telah ditabelkan untuk n = 3 sampai n = 15. Dengan sifat simetri distribusi itu terhadap $n \neq (n+1)/4$

kita peroleh $P\left(T^{+} \geq x\right)=P\left(T^{+} \leq x^{*}\right)$ dengan $x^{*}=\frac{n(n+1)}{4}-x.$ Nilai-nilai x dan x^{*} dalam tabel distribusi T^{+} memenuhi hubungan ini. Untuk melukiskan bagaimana menggunakan tabel distribusi T^{+} ini, sekali lagi kita kembali ke contoh data kilometer dalam Contoh 7.7. Di sana n = 12 dan nilai T^{+} yang kita amati sama dengan 62. Dari tabel kita peroleh $P\left(T^{+} \geq 61\right)=0{,}046.$ Jadi hipotesis nol ditolak pada tingkat signifikansi $\alpha=0{,}046$, dan ini diartikan ada kenaikan jarak kilometer yang signifikan dengan menggunakan jenis busi yang baru.

Dengan bertambah besarnya ukuran sampel n, distribusi T^{+} di bawah H_0 semakin mendekati distribusi normal dengan mean = $n(n+1)/4$ dan variansi = $n(n+1)(2n+1)/24$.

Pendekatan sampel besar untuk statistik peringkat bertanda:
Di bawah hipotesis nol

$$Z=\frac{T^{+}-n(n+1)/4}{\sqrt{n(n+1)(2n+1) / 24}}$$

Mendekati distribusi N(0; 1).

Kenyataan di atas dapat digunakan untuk melakukan uji peringkat bertanda bagi sampel-sampel besar.

Dalam menghitung statistik peringkat bertanda, nilai sama dapat terjadi dalam dua hal, yaitu pertama, beberapa selisih D_i mungkin nol (karena observasi X_1 sama dengan observasi X_2); dan kedua, nilai-nilai selisih D_i yang tidak nol mungkin mempunyai nilai mutlak yang sama. Nilai sama jenis pertama ditangani dengan membuang nilai-nilai nol dan mengubah ukuran sampel menjadi $n'=n$ - banyak nilai-nilai nol.

Nilai sama jenis kedua ditangani dengan memberi peringkat rata-rata pada setiap observasi dalam kelompok observasi yang sama itu.

Dengan sampel besar dapat digunakan pendekatan normal dengan mean

$$T^{+}=\frac{n'(n'+1)}{4}$$

$$\text{Variansi } T^{+} = \frac{n'(n'+1)(2n'+1)}{24} - \frac{1}{48}\sum_{j=1}^{\ell} a_j(a_j^2 - 1) \text{ dengan}$$

ℓ = banyak observasi sama

a_j = banyak elemen dalam observasi sama ke j, j=1, 2, …, ℓ

Contoh 7.8

Dalam suatu program penelitian makanan dipilih 20 sukarelawan untuk menilai dua jenis makanan dalam skala 0 – 100. Tabel 7.5 berikut menyajikan skor-skor itu dan selisihnya. Gunakan uji peringkat bertanda untuk menentukan apakah ada perbedaan yang signifikan antara skor kedua makanan itu.

Tabel 7.5
Skor Dua Jenis Makanan oleh 20 Orang.

Makanan A	70	85	73	75	65	50	80	71	80	51
Makanan B	65	41	45	80	84	50	71	52	42	78
Selisih	5	44	28	-5	-19	0	9	19	38	-27

Makanan A	72	76	79	65	59	72	84	90	56	57
Makanan B	62	38	80	65	54	67	87	90	38	43
Selisih	10	38	-1	0	5	5	-3	0	18	14

Nilai-nilai numerik berurut selisih-selisih yang tidak nol, peringkat-peringkatnya, dan serta tanda-tanda aljabarnya adalah.

Nilai Mutlak berurut dengan Tandanya	−1	-3	$\underbrace{5 \quad -5 \quad 5 \quad 5}$				9	10	14	18
Peringkatnya	1	2	4,5	4,5	4,5	4,5	7	8	9	10

Nilai Mutlak Berurut dengan Tandanya	$\underbrace{-19 \quad 19}$		-27	28	$\underbrace{38 \quad 38}$		44
Peringkatnya	11,5	11,5	13	14	15,5	15,5	17

Banyak selisih yang tidak nol adalah $n' = 17$ dan 3 nilai (mutlak) yang sama. Maka, statistik peringkat bertanda adalah:

T^{+} = 4,5 + 4,5 + 4,5 + 7 + 8 + 9 + 10 + 11,5 + 14 + 15,5 + 15,5 + 17

T^{+} =121

Untuk pendekatan normal sampel besar, mean variansi, dan deviasi standarnya adalah:

$$\text{Mean} = 17 \times \frac{18}{4} = 76,5$$

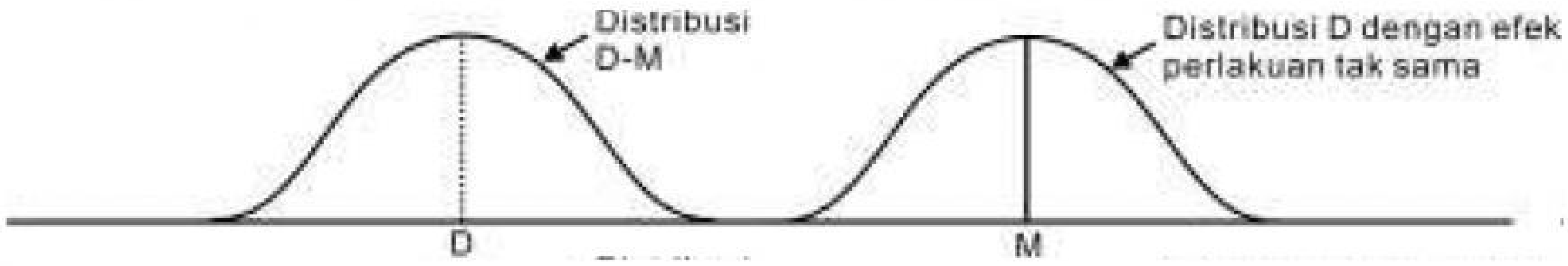
$$\text{variansi} = \frac{17 \times 18 \times 35}{24} - \frac{1}{48}(4 \times 15 + 2 \times 3 + 2 \times 3)$$
$$= 446,25 - 1,5 = 444,75$$

$$\text{deviasi standar} = \sqrt{444,75} = 21,09$$

Nilai standar statistik penguji itu adalah Z = (121 – 76,5)/21,09 = 2,11. Oleh karena nilai ini lebih besar dari nilai kritis normal standar dua sisi 1,96 yang berkaitan dengan tingkat signifikan 5% maka hipotesis nol tidak ada perbedaan dalam kualitas ditolak dengan statistik uji peringkat bertanda pada $\alpha = 0,05$.

3. Interval Kepercayaan berdasarkan Selisih Berpasangan

Kita perhatikan observasi berpasangan, seperti yang kita bicarakan di atas. Di bawah hipotesis nol bahwa kedua perlakuan itu sama maka selisih berpasangan $D_1, D_2, ..., D_n$ dianggap berdistribusi yang simetrik terhadap nol. Jika ada perbedaan dalam perlakuan, distribusi itu masih dianggap simetrik, tetapi mempunyai pusat simetrinya M. Oleh karena pusat ini membagi populasinya menjadi dua maka pusat ini juga merupakan median populasinya. Distribusi ini dilukiskan dalam Gambar 7.4 berikut.



Gambar 7.4
Distribusi Suatu Selisih

Statistik penguji peringkat bertanda Wilcoxon menguji $H_0 : M = 0$. Guna menguji hipotesis nol yang lebih umum $H_0 : M = M_0$. hanya diperlukan perubahan yang kecil saja. Di bawah hipotesis nol yang umum ini, kuantitas $(D_1 - M_0),...,(D_n - M_0)$ berdistribusi simetrik terhadap nol. Uji peringkat bertanda yang dilakukan dengan selisih terubah ini $D_i' = D_i - M_0$ merupakan uji untuk $H_0 : M = M_0$. Selanjutnya interval kepercayaan $(1-\alpha)100\%$ untuk M dapat dikonsepsikan sebagai nilai-nilai rentang M_0 yang menyebabkan hipotesis nol $H_0 : M = M_0$ tidak ditolak pada tingkat signifikansi α. Di sini kita tidak perlu melakukan uji peringkat bertanda dua sisi untuk semua nilai-nilai M_0 yang mungkin, melainkan kita lakukan langkah-langkah sebagai berikut:

a. Hitung nilai rata-rata semua selisih berpasangan, termasuk rata-rata $\frac{D_1 + D_1}{2},...,\frac{D_n + D_n}{2}$. Jadi, kita hitung

$$\frac{D_1 + D_1}{2},...,\frac{D_1 + D_n}{2},\frac{D_2 + D_2}{2},...,\frac{D_2 + D_n}{2},...,\frac{D_n + D_n}{2}$$

b. Urutkan $\binom{n}{2} + n = n(n+1)/2$ nilai rata-rata ini dari kecil ke besar.

c. Interval kepercayaan untuk pusat simetri M adalah interval terkecil ke d sampai terbesar ke d termasuk titik-titik ujungnya (nilai-nilai d dan koefisien kepercayaannya sudah ada dalam tabel).

Sebenarnya tidak perlu menghitung semua $n(n+1)/2$ nilai rata-rata. Dengan pengalaman yang kita punyai, kita dapat menentukan secara sistematis nilai rata-rata di kedua ujungnya setelah menyusun selisih D_i menurut besarnya dari kecil ke besar.

Contoh 7.9

Hitunglah interval kepercayaan untuk M berdasarkan data kilometer dalam contoh 7.7.

Menurut tabel yang tersedia (yakni tabel untuk menghitung interval kepercayaan bagi median menggunakan uji peringkat bertanda Wilcoxon) diperoleh d = 14 untuk suatu interval kepercayaan 95,8% dengan ukuran sampel n = 12. Jadi, hanya 14 nilai rata-rata terendah dan 14 nilai rata-rata tertinggi untuk selisih-selisih dalam Tabel 7.4. yang diperlukan. Duabelas nilai selisih itu dalam urutan dari kecil ke besar adalah.

- 1,1 -0,7 -0,3 -0,2 0,4 0,5 0,8 1,0 1,5 1,8 2,0 2,1

Rata-rata yang terkecil adalah [(-1,1)+(-1,1)]/2=-1,1, dan rata-rata berikutnya adalah [(-1,1)+(-0,7)]/2=-0,9. Seterusnya dengan cara seperti itu diperoleh:

Terkecil	-1,1	-0,9	-0,7	-0,7	-0,65	-0,5	-0,45	-0,35
Terbesar	1,4	1,45	1,5	1,5	1,55	1,65	1,75	1,8

Terkecil	-0,3	-0,3	-0,25	-0,2	-0,15	-0,15	...
Terbesar	1,8	1,9	1,95	2,0	2,05	2,1	

Maka interval kepercayaan 95,8% untuk M adalah [-0,15; 1,4].

Perhatikan bahwa nilai M = 0 termasuk di dalam interval itu, yang menunjukkan bahwa tidak ada perbedaan perlakuan merupakan kesimpulan yang dapat diterima.

Perlu dicatat bahwa dalam pengambilan kesimpulan ini, selisih nol harus tidak dihilangkan dan ukuran sampel harus tidak dikurangi seperti dalam pengujian. Jika ada hasil hitung yang sama, interval-interval itu setidaknya mempunyai tingkat kepercayaan seperti yang dinyatakan jika titik-titik tepi dimasukkan dalam interval. Jika diagram titik memperlihatkan kecenderungan yang sangat tidak simetrik, interval itu tidak cocok, dan hasil serupa berdasarkan uji tanda dapat digunakan untuk menentukan interval bagi median.

B. UKURAN KORELASI BERDASARKAN PERINGKAT

Peringkat dapat juga dimanfaatkan guna menentukan tingkat hubungan (keterkaitan antara dua variabel random, misalnya variabel tingkat kemampuan dalam matematika dan variabel tingkat apresiasi terhadap musik. Skor reaksi terhadap uji psikologis bagi anak kembar kita pelajari ukuran keterkaitan antara X dan Y dengan rumus

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}}$$

yang kita kenal sebagai koefisien korelasi hasil kali *Moment Pearson*. Sebagai statistik deskriptif, ukuran ini memberikan nilai numerik untuk besar keterkaitan linear antara X dan Y. Dalam hal ini distribusi sampling untuk r hanya berlaku di bawah anggapan bahwa distribusi bersama X dan Y normal. Metode korelasi peringkat tidak memerlukan anggapan normal dan bahkan menunjukkan sifat yang lebih kuat, yakni mengukur juga hubungan yang tidak linear.

Struktur observasi.
n pasangan $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ adalah independen dan tiap pasang mempunyai distribusi bivariat kontinu yang sama. Selanjutnya $X_1, X_2, ..., X_n$ diberi peringkat, demikian juga $Y_1, Y_2, ..., Y_n$

Pasangan	1	2	...	n
Peringkat X	R_1	R_2	...	R_n
Peringkat Y	S_1	S_2	...	S_n

Kita catat beberapa sifat yang menyederhanakan hitungan $R_1, R_2, ..., R_n$ adalah angka-angka peringkat Y. Jadi, merupakan satu permutasi (urutan) angka-angka 1, 2, …, n. Oleh karena itu, nilai rata-ratanya adalah $\overline{R} = (1+2+...+n)/n =$ $[n(n+1)]/2n = (n+1)/2$. Demikian juga $\overline{S}$ selalu sama dengan $(n+1)/2$. Selanjutnya dapat ditunjukkan

$$\sum_{i=1}^{n}\left(R_i - \overline{R}\right)^2 = \frac{n\left(n^2-1\right)}{12}$$ dan juga

$$\sum_{i=1}^{n}\left(S_i - \overline{S}\right)^2 = \frac{n\left(n^2-1\right)}{12}$$

Suatu ukuran korelasi didefinisikan oleh C. Spearman yang analog dengan koefisien korelasinya Pearson. Spearman mengganti nilai observasi dengan peringkatnya. Koefisien korelasi peringkat Spearman didefinisikan dengan

$$r_{Sp} = \frac{\sum_{i=1}^{n}(R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{n}(R_i - \bar{R})^2}\sqrt{\sum_{i=1}^{n}(S_i - \bar{S})^2}}$$
$$= \frac{\sum_{i=1}^{n}\left(R_i - \frac{n+1}{2}\right)\left(S_i - \frac{n+1}{2}\right)}{n(n^2-1)/12}$$

Koefisien korelasi Spearman ini juga memenuhi $-1 \le r_{Sp} \le 1,$ dan nilai-nilai dekat +1 menunjukkan kecenderungan nilai-nilai X yang lebih besar berpasangan dengan nilai-nilai yang lebih besar pula. Tetapi, korelasi peringkat lebih berarti karena interpretasinya tidak harus mengacu hubungan yang linear.

Korelasi Peringkat Spearman:

$$r_{Sp} = \frac{\sum_{i=1}^{n}\left(R_i - \frac{n+1}{2}\right)\left(S_i - \frac{n+1}{2}\right)}{n\left(n^2-1\right)/12}$$

a. $-1 \le r_{Sp} \le 1,$

b. r_{Sp} dekat dengan +1 menunjukkan suatu kecenderungan nilai-nilai X yang lebih besar berkaitan dengan nilai-nilai Y yang lebih besar pula. Nilai r_{Sp} dengan –1 menunjukkan hubungan yang sebaliknya.

c. Hubungan itu tidak perlu linear; hanya hubungan naik/turun yang diperlukan.

Contoh 7.10:

Ingin dipelajari kuat hubungan antara peringkat hasil *interview* dengan nilai kerja enam karyawan baru di suatu perusahaan.

Interview (Peringkat)	5	2	3	1	6	4
Nilai Kerja	47	32	29	28	56	38

Jika nilai kerja itu kita urutkan dan kita peringkat, kita peroleh

Interview (R_i)	5	2	3	1	6	4
Nilai kerja (S_i)	5	3	2	1	6	4

Selanjutnya akan kita hitung r_{Sp}.

$$\overline{R} = \frac{(n+1)}{2} = \frac{(6+1)}{2} = 3,5$$

$$n\frac{(n^2-1)}{12} = 6\frac{(36-1)}{12} = \frac{35}{2} = 17,5$$

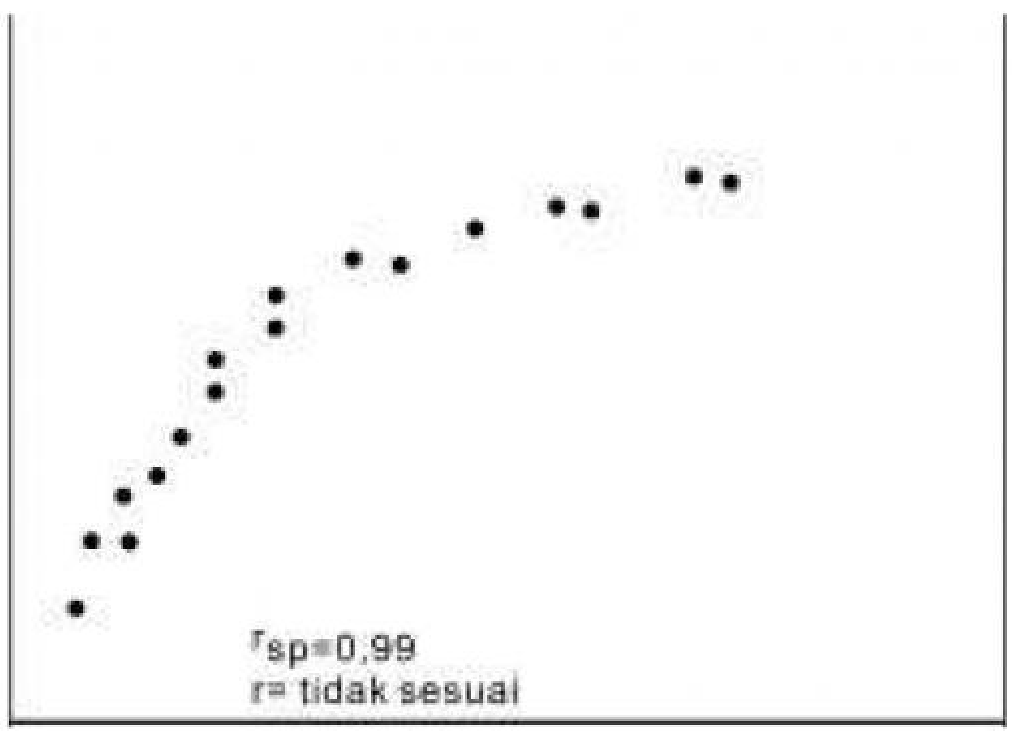
maka

$$\sum_{i=1}^{n}\left(R_i - \frac{n+1}{2}\right)\left(S_i - \frac{n+1}{2}\right)$$
$$= (5-3,5)(5-3,5) + (2-3,5)(3-3,5) + ... + (4-3,5)(4-3,5)$$
$$= (1,5)(1,5) + (-1,5)(-0,5) + ... + (0,5)(0,5)$$
$$= 16,5$$

dan

$$r_{Sp} = \frac{16,5}{17,5} = 0,943$$

Gambar 7.5 berikut ini dapat membantu menegaskan bahwa r_{Sp} adalah ukuran untuk sebarang hubungan yang monoton, bukan hanya hubungan yang linear.



Gambar 7.5
r_{sp} adalah Ukuran Hubungan Monoton

Tanpa mengemukakan lagi anggapan tentang populasinya, kita dapat menggunakan korelasi peringkat untuk menguji independensi X dan Y versus alternatif hubungan yang naik (turun), tetapi sebelum kita melakukan hal ini, kita akan melihat lebih jauh bagaimana menyederhanakan hitungan statistik ini. Karena

$$\sum_{i=1}^{n}\left(R_i-\frac{n+1}{2}\right)\left(S_i-\frac{n+1}{2}\right)$$
$$=\sum_{i=1}^{n}R_iS_i-\frac{n+1}{2}\sum_{i=1}^{n}R_i-\frac{n+1}{2}\sum_{i=1}^{n}S_i+n\left(\frac{n+1}{2}\right)^2$$
$$=\sum_{i=1}^{n}R_iS_i-n\left(\frac{n+1}{2}\right)^2.$$

maka bagian yang random dari r_{Sp} hanyalah $\sum_{i=1}^{n}R_iS_i$. Dengan demikian, kita dapat mendasarkan uji kita pada kuantitas ini dan mentabelkan distribusinya, setidaknya untuk n kecil.

Hipotesis untuk Uji Independensi

$(X_1,Y_1),(X_2,Y_2),...,(X_n,Y_n)$ adalah sampel random dari suatu distribusi bivariat kontinu.

$H_0 : X$ dan Y independen, yakni

$\mathrm{P}(\mathrm{X}\le \mathrm{x};\mathrm{Y}\le \mathrm{y})=\mathrm{P}(\mathrm{X}\le \mathrm{x})\mathrm{P}(\mathrm{Y}\le \mathrm{y})$

untuk semua x dan y.

Sebagai akibat dari H_0 adalah semua himpunan n! pasangan (R_i,S_i) berkemungkinan sama, masing-masing berpeluang $1/\mathrm{n}!$.

Alternatif satu sisi

H_1 : nilai-nilai X yang besar mempunyai kecenderungan akan terjadi dengan nilai-nilai yang besar.

Alternatif dua sisi

H_1 : (a) nilai-nilai X yang besar mempunyai kecenderungan akan terjadi dengan nilai-nilai Y yang besar; atau

(b) nilai-nilai X yang kecil mempunyai kecenderungan akan terjadi dengan nilai-nilai Y yang besar.

Uji yang didasarkan pada $\sum_{i=1}^{n} R_i S_i$ atau secara ekuivalen, berdasarkan r_{Sp}, dilakukan dengan menggunakan distribusi $\sum_{i=1}^{n} R_i S_i$ yang telah ditabelkan untuk $n \leq 10$.

Uji Independensi Berdasarkan r_{Sp}

Alternatif satu sisi

H_0 ditolak jika $T = \sum_{i=1}^{n} R_i S_i \geq x$

dengan

$$P\left[\sum_{i=1}^{n} R_i S_i \geq x\right] \leq \alpha$$

Alternatif dua sisi

H_0 ditolak jika $\sum_{i=1}^{n} R_i S_i \geq x$ *atau* $\leq x^*$

dengan

$$\frac{\alpha}{2} = P\left[\sum_{i=1}^{n} R_i S_i \geq x\right] = P\left[\sum_{i=1}^{n} R_i S_i \leq x^*\right]$$

Dalam sampel-sampel besar dengan n lebih besar dari yang ada dalam tabel distribusi $\sum_{i=1}^{n} R_i S_i$ kita dapat menggunakan kenyataan berikut.

Di bawah $H_0, \sqrt{n-1}\, r_{Sp}$ mendekati distribusi N(0; 1)

Dengan perkataan lain, dengan H_1 satu sisi, H_0 ditolak dalam sampel-sampel besar jika $\sqrt{n-1}\ r_{Sp} > Z_\alpha$, yakni titik yang membatasi luasan α di atas distribusi normal standar.

Contoh 7.11.

Kembali ke Contoh 7.10, ujilah independensi dari peringkat hasil interviu dan nilai kerja. Kita pandang alternatif dua sisi, yakni korelasi positif atau negatif. Nilai statistik penguji yang diamati adalah

$$\sum_{i=1}^{6} R_i S_i = 5.5 + 2.3 + 3.2 + 1.1 + 6.6 + 4.4 = 90$$

Dari tabel distribusi $\sum_{i=1}^{n} R_i S_i$, kita lihat bahwa nilai kritis ekor atas 89 mempunyai peluang $P\left(\sum R_i S_i \geq 89\right) = 0,017$ sehingga nilai $\sum_{i=1}^{6} R_i S_i = 89$ yang diamati jatuh pada daerah kritis untuk $\alpha = 0,05$ dan $\frac{\alpha}{2} = 0,025$ untuk setiap ekor. Maka, hipotesis independensi ditolak.

Jika ada observasi yang sama, pendekatan normal sampel menjadi:

$$\frac{\sqrt{n-1}\ r_{Sp}}{\sqrt{\left(1-C_x\right)\left(1-C_y\right)}} \text{ mendekati N(0; 1)}$$

dengan

$$C_x = \sum_{j=1}^{\ell} \frac{q_j^3 - q_j}{n^3 - n}$$

q_j = banyak elemen sama ke j sampel X

dan

C_y = koreksi serupa berkaitan dengan sampel Y.

Koefisien tau Kendall

Satu korelasi peringkat yang juga sering digunakan adalah koefisien Tau Kendall yang didefinisikan sebagai:

$$r_{\text{tau}} = \sum_{\substack{\textit{semua pasangan} \\ (i,j),\ i<j}} \frac{\text{tanda}\left(S_i - S_j\right)\text{tanda}\left(R_i - R_j\right)}{\frac{1}{2}n\left(n-1\right)}$$

dengan tanda $\left(R_i - R_j\right)$=+1 jika $R_i > R_j$ dan -1 jika $R_i < R_j$.

Tau Kendall mempunyai sifat-sifat yang sama seperti r_{Sp} dan dalam banyak hal kedua korelasi itu serupa. Tabel distribusi dan informasi tentang distribusi sampel besar tan Kendall telah banyak dipelajari orang, tetapi tidak kita pelajari di sini.

Dua pasangan $\left(X_1, Y_1\right)$ dan $\left(X_2, Y_2\right)$ dikatakan *sesuai jika* $\left(X_1 - Y_1\right)$ dan $\left(X_2 - Y_2\right)$ mempunyai tanda yang sama jika tidak demikian dikatakan *tak sesuai*. Korelasi r_{tau} adalah penaksir tak bias untuk

$$\begin{aligned} P(sesuai)\ \text{-}\ P(tak\ sesuai) &= 2\ P(sesuai) - 1 \\ &= 2\ P\left[\left(X_1 - X_2\right)\left(Y_1 - Y_2\right) > 0\right] - 1 \end{aligned}$$

Di bawah hipotesis independensi, nilai harapan r_{tau} sama dengan nol.

LATIHAN

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

1) Seorang peneliti sosial melakukan interview terhadap 25 pasangan yang baru saja menikah. Suami dan istri masing-masing secara independen ditanya "Berapa banyak anak yang diinginkan?" Diperoleh data sebagai berikut.

Pasangan	Jawaban	
	Suami	Istri
1	3	2
2	1	1
3	2	1

Pasangan	Jawaban	
	Suami	Istri
14	2	1
15	3	2
16	2	2

Pasangan	Jawaban	
	Suami	Istri
4	2	3
5	5	1
6	0	1
7	0	2
8	1	3
9	2	2
10	3	1
11	4	2
12	1	2
13	3	3

Pasangan	Jawaban	
	Suami	Istri
17	0	0
18	1	2
19	2	1
20	3	2
21	4	3
22	3	1
23	0	0
24	1	2
25	1	1

Apakah data menunjukkan perbedaan pandangan yang signifikan antara suami dan istri tentang keluarga yang ideal? Gunakan uji tanda dengan tingkat signifikansi mendekati 5%.

2) Distribusi statistik peringkat bertanda Wilcoxon T^+ ditentukan dari kenyataan bahwa di bawah hipotesis nol distribusinya simetrik terhadap nol, masing-masing peringkat 1, 2, …, n berkemungkinan sama akan berkaitan dengan tanda positif atau negatif. Lagipula, tanda-tanda itu independen dengan peringkat:
   a. dengan memandang kasus $n=3$, identifikasikan semua $2^3=8$ perkaitan yang mungkin antara tanda dan peringkat 1, 2, dan 3, dan tentukan nilai T^+ untuk setiap perkaitan,
   b. berilah peluang yang sama dengan $\frac{1}{8}$ untuk setiap kasus, dapatkan distribusi T^+ dan periksalah bahwa peluang-peluang ekornya sama dengan yang sudah ditabelkan.

3) Skor berikut diperoleh dari uji keterampilan dan kegiatan yang diperoleh dari suatu sampel random 10 orang siswa sekolah menengah.

Siswa	1	2	3	4	5	6	7	8	9	10
Keterampilan	23	29	45	36	49	41	30	15	42	39
Kegiatan	45	48	16	28	38	21	36	18	31	37

Dengan menggunakan statistik Spearman, ujilah hipotesis nol bahwa manifestasi keterampilan dan kegiatan adalah independen!

4) Seorang peneliti melakukan percobaan guna menentukan apakah pupuk A dan pupuk B memberikan laju pertumbuhan yang berbeda. Pasangan tanaman tertentu, satu diberi pupuk A dan yang lain diberi pupuk B ditanam dalam pot, dan tingginya diukur setelah periode waktu tertentu, diperoleh data sebagai berikut.

Pasangan	Tinggi Tanaman (cm)	
	Pupuk A	Pupuk B
1	188	139
2	96	163
3	168	160
4	176	160
5	153	147
6	172	149
7	177	149
8	163	122
9	146	132
10	173	144
11	186	130
12	168	144
13	177	102
14	184	124
15	96	144

a. Hitunglah selisih berpasangan dan gambarkan diagram titik untuk data itu. Apakah anggapan distribusi normal dapat diterima?
b. Lakukanlah uji peringkat bertanda Wilcoxon untuk menentukan apakah pupuk A memberikan tingkat pertumbuhan yang lebih besar daripada pupuk B
c. Hitunglah interval kepercayaan 95% untuk selisih tingkat pertumbuhan Δ.

RANGKUMAN

Kita pelajari dua uji nonparametrik untuk data berpasangan, yakni uji tanda dan uji peringkat bertanda Wilcoxon, baik untuk sampel kecil maupun besar.

Kita pelajari juga bagaimana menghitung interval kepercayaan $(1-\alpha)100\%$ untuk pusat simetri M.

Selanjutnya kita pelajari dua koefisien korelasi peringkat, yakni koefisien korelasi Spearman dan Koefisien korelasi Tau Kendall. Kita akhiri dengan uji independensi.

TES FORMATIF 2

Pilihlah satu jawaban yang paling tepat!

I. Enam orang melakukan diet teratur dalam upaya mengurangi berat badan. Diperoleh hasil pengukuran berat badan sebelum dan sesudah diet sebagai berikut (dalam kg)

Orang	1	2	3	4	5	6
Sebelum	80	88	87	84	93	87
Sesudah	76	86	84	82	94	83

1) Untuk uji tanda dengan $H_0 : \mathrm{P}(+) = 0{,}5$ kita peroleh statistik penguji S sama dengan ....
A. 0,5
B. 1,0
C. 1,5
D. 2,0

2) Dengan daerah kritis $P(S \geq 1)$ kita punyai tingkat signifikasi α sama dengan ....
A. 0,4950
B. 0,3780
C. 0,2650
D. 0,1094

3) Dengan uji peringkat bertanda Wilcoxon kita peroleh statistik penguji T^+ sama dengan ....
A. $T^+ = 20$
B. $T^+ = 17$
C. $T^+ = 10$
D. $T^+ = 8$

4) Interval kepercayaan (kira-kira) 95% untuk pusat simetri M adalah ....
   A. [0,4; 4,5]
   B. [0,4; 3,8]
   C. [0,5; 4,5]
   D. [0,6; 3,5]

II. Duapuluh lima orang siswa SMU disuatu kota diberi ujian verbal dan ujian kuantitatif. Diperoleh data sebagai berikut.

Siswa	1	2	3	4	5	6	7	8	9	10	11	12	13
Ujian verbal	41	32	68	87	75	53	94	18	36	49	63	55	43
Ujian kuantitatif	53	41	63	90	60	62	82	29	39	36	58	67	40

Siswa	14	15	16	17	18	19	20	21	22	23	24	25
Ujian verbal	48	61	71	52	49	29	68	37	57	84	42	77
Ujian kuantitatif	48	47	87	43	78	22	60	53	48	70	31	91

5) Untuk uji tanda bagi $H_0 : \mathrm{P}(+) = 0{,}5$ versus $H_1 : \mathrm{P}(+) \neq 0{,}5$ kita peroleh statistik penguji Z sama dengan ....
   A. 0,37
   B. 0,41
   C. 0,56
   D. 0,64

6) Untuk uji dalam 5), dengan tingkat signifikansi $\alpha = 5\%$ kita peroleh daerah kritik ....
   A. $Z < 1{,}96$ atau $> 1{,}96$
   B. $Z < 1{,}64$ atau $> 1{,}64$
   C. $Z < -1{,}96$ atau $> 1{,}96$
   D. $Z < -1{,}64$ atau $> 1{,}64$

7) Untuk uji peringkat bertanda Wilcoxon kita peroleh Mean T^+ sama dengan ....
   A. 150
   B. 159
   C. 162
   D. 168

8) Variansi T^+ sama dengan ....
   A. 934,255
   B. 987,635
   C. 1024,365
   D. 1223,875

9) Statistik penguji Z sama dengan ....
   A. 0,28
   B. 0,36
   C. 0,47
   D. 0,59

10) Koefisien korelasi Spearman sama dengan ....
   A. 0,777
   B. 0,812
   C. 0,888
   D. 0,912

11) Untuk uji independensi kita peroleh statistik penguji sampel besar sama dengan ….
   A. 2,03
   B. 2,35
   C. 2,87
   D. 3,81

Cocokkanlah jawaban Anda dengan Kunci Jawaban Tes Formatif 2 yang terdapat di bagian akhir modul ini. Hitunglah jawaban yang benar. Kemudian, gunakan rumus berikut untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 2.

$$\text{Tingkat penguasaan} = \frac{\text{Jumlah Jawaban yang Benar}}{\text{Jumlah Soal}} \times 100\%$$

Arti tingkat penguasaan: 90 - 100% = baik sekali
80 - 89% = baik
70 - 79% = cukup
< 70% = kurang

Apabila mencapai tingkat penguasaan 80% atau lebih, Anda dapat meneruskan dengan modul berikutnya. **Bagus!** Jika masih di bawah 80%, Anda harus mengulangi materi Kegiatan Belajar 2, terutama bagian yang belum dikuasai.

Kunci Jawaban Tes Formatif

Tes Formatif 1		*Tes Formatif 2*	
1)	B	1)	B
2)	A	2)	D
3)	D	3)	A
4)	B	4)	C
5)	A	5)	B
6)	D	6)	C
7)	C	7)	A
8)	B	8)	D
9)	A	9)	B
10)	C	10)	A
		11)	D

Daftar Pustaka

Battacharyya, G. K. and R.A. Johnson (1977). *Statistics Concepts and Methods*. New York: John Willey.

Freud J. (1979). *Modern Elementary Statistics*. Prentice Hall.

Kooros, A. (1965). *Elements of Mathematical Economics*. Boston: Houghton Mifflin Company.

Pfeffenberger, R.C. and J.H. Peterson (1977). *Statistical Methods for Business and Economics*. Illinois: Richard D. Irwin.

Lampiran

Table Quantiles of the Wilcoxon Signed Ranks test Statistic

	$W_{.005}$	W1	$W_{.025}$	$W_{.05}$	$W_{.10}$	$W_{.20}$	$W_{.30}$	$W_{.40}$	$W_{.50}$	$\frac{n(n+1)}{2}$
N=4	0	0	0	0	1	3	3	4	5	10
5	0	0	0	1	3	4	5	6	7,5	15
6	0	0	1	3	4	6	8	9	10,5	21
7	0	1	3	4	6	9	11	12	14	28
8	1	2	4	6	9	12	14	16	18	36
9	2	4	6	9	11	15	18	20	22,5	45
10	4	6	9	11	15	19	22	25	27,5	55
11	6	8	11	14	18	23	27	30	33	66
12	8	10	14	18	22	28	32	36	39	78
13	10	13	18	22	27	33	38	42	45,5	91
14	13	16	22	26	32	39	44	48	52,5	105
15	16	20	26	31	37	45	51	55	60	120
16	20	24	30	36	43	51	58	63	68	136
17	24	28	35	42	49	58	65	71	76,5	153
18	28	33	41	48	56	66	73	80	85,5	171
19	33	38	47	54	63	74	82	89	95	190
20	38	44	53	61	70	82	91	98	105	210