NEPARAMETARSKI TESTOVI

- Mnogi testovi pretpostavljaju normalnu razdiobu populacije/a.
- Ukoliko je ta pretpostavka narušena, koristimo testove koji nemaju pretpostavke o distribuciji populacije (**neparametarske testove**)
- Primjer neparametarskog testa je
  - **Pearsonov χ^2-test**
  - **Kolmogorov-Smirnov test**
- Neparametarski testovi su manje osjetljivi na ekstremne vrijednosti (outliere) od parametarskih.
- Neparametarski testovi uglavnom imaju manju snagu od parametarskih kada su ispunjene pretpostavke parametarskih testova.

Medijan test

Nul-hipotezu koja se testira je da dvije populacije imaju isti medijan.

Dvije populacije: A i B.

Testiramo hipotezu o jednakosti medijana tih dviju populacija:

$$H_0 : \quad M_A = M_B$$

Iz svake se populacije izabere po jedan uzorak (ne nužno iste veličine).

Na osnovu kombiniranog uzorka (nastalog spajanjem ova dva uzorka) odredi se kombinirani (zajednički) medijan.

Ukoliko je pretpostavka o jednakim medijanima točna, tada je udio jedinki ispod, odnosno iznad kombiniranog medijana otprilike podjednak u oba uzorka (grupe).

Daljnja analiza se svodi na proučavanje 2×2 tablice:

	Grupa		
	Uzorak 1	Uzorak 2	Kombinirano
Iznad komb. medijana	A	B	$A+B$
Ispod komb. medijana	C	D	$C+D$
Ukupno	$m\ (=A+C)$	$n\ (=B+D)$	

Za testiranje jednakosti distribucija koristi se Fisherov egzaktni test.

Metodu ćemo ilustrirati na primjeru korištenom za uspoedbu srednjih vrijednosti.

Primjer. U studiji o razlici u apsolutnoj pogrešci kod pozicioniranja aktivne i pasivne ruke istraživači su testirali 20 studenata. Kod 10 studenata su mjerili apsolutnu pogrešku kod pozicioniranja aktivne ruke a kod preostalih 10 pogrešku kod pozicioniranja pasivne ruke. Rezultati mjerenja (u cm) prikazani su u tablici:

Ispitanik	Aktivna ruka	Ispitanik	Pasivna ruka
1	2.65	11	3.30
2	2.42	12	2.00
3	3.30	13	0.09
4	0.19	14	0.04
5	1.25	15	4.56
6	2.00	16	3.33
7	3.34	17	1.02
8	4.08	18	0.89
9	0.70	19	2.78
10	2.89	20	1.65

Tražimo kombinirani medijan.

Sve podatke poredamo po veličini

0.04	0.09	0.19	0.7	0.89	1.02	1.25	1.65	2	2
2.42	2.65	2.78	2.89	3.3	3.3	3.33	3.34	4.08	4.56

$$\text{kombinirani medijan } = 2.21$$

Sada prebrojimo podatke iznad i ispod kombiniranog medijana.

kombinirani medijan $= 2.21$

	Aktivna ruka	Pasivna ruka
	2.65	3.30
	2.42	2.00
	3.30	0.09
	0.19	0.04
	1.25	4.56
	2.00	3.33
	3.34	1.02
	4.08	0.89
	0.70	2.78
	2.89	1.65
Iznad	6	4
Ispod	4	6

R

```
> podaci
       Aktivna pasivna
Iznad        6       4
Ispod        4       6
> fisher.test(podaci)

        Fisher's Exact Test for Count Data

data:  podaci
p-value = 0.6563
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  0.2773893 19.1425577
sample estimates:
odds ratio
  2.158166
```

Fisherov egzaktni dvostrani test: $p = 0.6563$

Wilcoxon-Mann-Whitneyev test

- Neparametarski analogon Studentovom t-testu za nezavisne populacije.
- U literaturi je poznat i pod nazivima
  - **Wilcoxonov test sume rangova**
  - **Mann-Whitneyev U-test**
- **Nul hipoteza.**
  - X - slučajno izabrana jedinka iz prve populacije
  - Y - slučajno izabrana jedinka iz druge populacije

  H_0: Distribucija obilježja je jednaka u obje populacije.
- **Alternativna hipoteza.**

  H_1: Medijani dviju populacija su različiti.
- Test može biti jednostran i dvostran.

Pretpostavke testa.

- Distribucija obilježja je neprekidno ili ordinalno.
- Uzorci su izabrani nezavisno.
- Standardne devijacije u obje populacije su jednake.

Test.

- Iz svake populacije nezavisno izaberemo po jedan uzorak (ne nužno iste veličine)
  - n_1 - veličina uzorka iz prve populacije
  - n_2 - veličina uzorka iz druge populacije
- Svakom podatku se pridijeli rang iz kombiniranog uzorka.

  Tako pridijeljeni rangovi poprimaju vrijednosti od 1 do $n_1 + n_2$.
- Ukoliko je obilježje jednako distribuirano unutar obje populacije tada se očekuje da će prosječan rang biti podjednak za oba uzorka.
- Statistika:

$$W = \text{zbroj rangova iz jednog od uzorka.}$$

- Koji će se uzorak izabrati nije važno jer je ukupan zbroj rangova za $N = n_1 + n_2$ opservacija jednak
$$1 + 2 + 3 + \ldots + N = \frac{N(N+1)}{2}.$$
Zbroj rangova za drugi uzorak je
$$\frac{N(N+1)}{2} - W.$$
- Vrijedi:
$$\begin{aligned} E(W) &= \frac{n_1(N+1)}{2} \\ Var(W) &= \frac{n_1 n_2 (N+1)}{12} \end{aligned}$$
- Za veliki uzorak je W približno distribuiran prema normalnoj distribuciji ($n_1 \geq 10$ ili $n_2 \geq 10$; $N \geq 20$).

- U literaturi se često koristi Mann-Whitneyev U-test.

  Test je ekvivalentan ovom testu jer statistika U u MW testu je

$$U = n_1 n_2 - \frac{n_1(n_1 + 1)}{2} - W$$

  Ovdje je korištena statistika korištena u varijanti poznatoj pod imenom Wilcoxonov test sume rangova.

Izjednačeni rangovi

- Wilcoxonov test pretpostavlja da je obilježje neprekidno distribuirano.
- S jako preciznim mjerenjem neprekidne varijable, vjerojatnost istih vrijednosti (time i rangova) u dva (ili više) mjerenja je 0.
- Relativno grubo mjerenje može uzrokovati izjednačene vrijednosti.
- U slučaju jednakih vrijednosti, svakoj opservaciji pridružujemo srednju vrijednost rangova koje bi imale da izjednačenja nema.
- Ukoliko se izjednačenje dogodi između dvije ili više vrijednosti unutar samo jedne grupe to nema utjecaja na statistiku W.
- Ukoliko se to dogodi na dvije ili više opservacija iz obje grupe pojavljuje se utjecaj na W
- Međutim, svaka pojava izjednačenja utječe na varijabilnost rangova, a time i na varijancu statistike W.

- Varijanca

$$Var(W) = \frac{n_1 n_2 (N+1)}{12}$$

izračunata je uz pretpostavku da nema izjednačenih rangova.

- U slučaju izjednačenja potrebno je koristiti korekcije.
    - g - broj različitih vrijednosti za koje su rangovi izjednačeni
    - t_j - broj opservacija s istom j-tom vrijednošću ($j = 1, \ldots, g$)

Korigirana varijanca:

$$Var(W) = \frac{n_1 n_2 (N+1)}{12} - \frac{n_1 n_2}{12N(N-1)} \sum_j \left(t_j^3 - t_j \right).$$

Primjer. U studiji o razlici u apsolutnoj pogrešci kod pozicioniranja aktivne i pasivne ruke istraživači su testirali 20 studenata. Kod 10 studenata su mjerili apsolutnu pogrešku kod pozicioniranja aktivne ruke a kod preostalih 10 pogrešku kod pozicioniranja pasivne ruke. Rezultati mjerenja (u cm) prikazani su u tablici:

Ispitanik	Aktivna ruka	Ispitanik	Pasivna ruka
1	2.65	11	3.30
2	2.42	12	2.00
3	3.30	13	0.09
4	0.19	14	0.04
5	1.25	15	4.56
6	2.00	16	3.33
7	3.34	17	1.02
8	4.08	18	0.89
9	0.70	19	2.78
10	2.89	20	1.65

Odredimo rangove kombiniranog uzorka.

0.04	0.09	0.19	0.7	0.89	1.02	1.25	1.65	2	2
1	2	3	4	5	6	7	8	9.5	9.5
2.42	2.65	2.78	2.89	3.3	3.3	3.33	3.34	4.08	4.56
11	12	13	14	15.5	15.5	17	18	19	20

	Aktivna ruka		Pasivna ruka	
	Pogreška	**Rang**	**Pogreška**	**Rang**
	2.65	12	3.30	15.5
	2.42	11	2.00	9.5
	3.30	15.5	0.09	2
	0.19	3	0.04	1
	1.25	7	4.56	20
	2.00	9.5	3.33	17
	3.34	18	1.02	6
	4.08	19	0.89	5
	0.70	4	2.78	13
	2.89	14	1.65	8
W		**113**		(97)

$$\begin{aligned} n_1 &= 10 \\ n_2 &= 10 \\ N &= n_1 + n_2 = 20 \end{aligned}$$

R

R Console

```
> grupa
 [1] 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2
> pogreska
 [1] 2.65 2.42 3.30 0.19 1.25 2.00 3.34 4.08 0.70 2.89 3.30 2.00
[13] 0.09 0.04 4.56 3.33 1.02 0.89 2.78 1.65
> wilcox.test(pogreska ~ grupa, data=)

        Wilcoxon rank sum test with continuity correction

data:  pogreska by grupa
W = 58, p-value = 0.5705
alternative hypothesis: true location shift is not equal to 0

Warning message:
In wilcox.test.default(x = c(2.65, 2.42, 3.3, 0.19, 1.25, 2, 3.34,  :
  cannot compute exact p-value with ties
> 
```

- W - statistika ($= W - 10 \cdot (10+1)/2 = W - 55$), p-value - normalna aproksimacija s korekcijom za izjednačene rangove
- $p = 0.5705 \quad \Rightarrow \quad$ nema razlike između aktivne i pasivne ruke

Kruskal-Wallisov test

- Neparametarski analogon ANOVA testu s jednostrukom klasifikacijom.
- U literaturi je poznat i pod nazivima
  - Kruskal-Wallis ANOVA
  - neparametarska ANOVA
- Kao i Wilcoxon-Mann-Witneyev test, Kruskal-Wallisov test koristi rangove, te su nazivi koji sadrže ANOVA loši.
- Kruskal-Wallisov test uspoređuje srednje rangove koristeći usporedbu varijanci rangova na isti način kao što je to napravljeno u ANOVA testu.
- KW test je proširenje Wilcoxon-Mann-Witneyevog testa na usporedbu više nezavisnih populacija.

- **Nul hipoteza.**

  H_0: Distribucija obilježja je jednaka u svim populacijama.

- **Alternativna hipoteza.**

  H_1: Medijani barem dvije populacije su različiti.

Pretpostavke testa.

- Distribucija obilježja je neprekidno ili ordinalno.
- Uzorci su izabrani nezavisno.
- Standardne devijacije u svim populacijama su jednake.

Test.

- Iz svake od k populacija nezavisno izaberemo po jedan uzorak (ne nužno iste veličine):
  n_i - veličina uzorka iz i-te populacije ($i = 1, \ldots, k$)
- Svakom podatku se pridijeli rang iz kombiniranog uzorka.

  Tako pridijeljeni rangovi poprimaju vrijednosti od 1 do $N = n_1 + n_2 + \ldots + n_k$.
- Ukoliko je obilježje jednako distribuirano unutar obje populacije tada se očekuje da će prosječan rang biti podjednak za sve uzorka.

Podaci. Izračunamo rangove na kombiniranom uzorku.

Grupa			
1	**2**	**...**	**k**
X_{11}	X_{21}	$\dots$	X_{k1}
X_{12}	X_{22}	$\dots$	X_{k2}
$\vdots$	$\vdots$		$\vdots$
X_{1n_1}	$\vdots$		$\vdots$
	$\vdots$		X_{2n_k}
	X_{2n_2}		

Grupa			
1	**2**	**...**	**k**
R_{11}	R_{21}	$\dots$	R_{k1}
R_{12}	R_{22}	$\dots$	R_{k2}
$\vdots$	$\vdots$		$\vdots$
R_{1n_1}	$\vdots$		$\vdots$
	$\vdots$		R_{2n_k}
	R_{2n_2}		
$\bar{R}_1$	$\bar{R}_2$	$\dots$	$\bar{R}_k$

$\bar{R}_i$ srednja vrijednost rangova u i-tom uzorku

Statistika.

$$KW = \frac{12}{N(N+1)} \sum_i n_i \left(\bar{R}_i - \bar{R}\right)^2$$

n_i - veličina i-tog uzorka

N - veličina kombiniranog uzorka ($N = n_1 + n_2 + \ldots + n_k$)

$\bar{R}_i$ - srednja vrijednost rangova u i-tom uzorku

$\bar{R}$ - srednja vrijednost rangova u kombiniranom uzorku.

Ukoliko je broj grupa veći od 3 ($k \geq 3$) i veličina svakog uzorka veća od 5 ($n_i \geq 5$), tada je razdioba statistike KW dobro aproksimirana χ^2 razdiobom s $k-1$ stupnjem slobode.

Izjednačeni rangovi

- Kao i kod Wilcoxonovog testa, pretpostavlja se da je obilježje neprekidno distribuirano te je vjerojatnost istih vrijednosti u dva (ili više) mjerenja 0.
- Ukoliko postoje izjednačeni rangovi, mjenja se varijanca rangova te treba korigirati *KW* statistiku.
- *KW* statistika se dijeli s

$$1 - \frac{1}{N^3 - N} \sum_{j=1}^{g} \left(t_j^3 - t_j \right).$$

  - g - broj različitih vrijednosti za koje su rangovi izjednačeni
  - t_j - broj opservacija s istom j-tom vrijednošću ($j = 1, \ldots, g$)

’Post-hoc’ test

Kao i ANOVA, Kruskal-Wallisov test u slučaju odbacivanja hipoteze pokazuje da se populacije razlikuju ali ne daje odgovor na pitanje koje se populacije razlikuju.

’Post-hoc’ test: **Dunnov test**

Primjer. Trener želi usporediti tri različite metode treninga. Svaku od metoda primijenio je na po $n = 4$ studenta. Nakon 30 dana ocijenjena je uspješnost i ocjene su prikazane u tablici

Metoda	**Opservacije**			
Metoda 1	3	6	4	7
Metoda 2	11	8	10	7
Metoda 3	6	9	5	8

Jesu li sve tri metode jednako uspješne? Hipotezu testirajte uz razinu značajnosti $\alpha = 0.05$.

Rangiranje podataka.

3	4	5	6	6	7	7	8	8	9	10	11
1	2	3	4.5	4.5	6.5	6.5	8.5	8.5	10	11	12

1. uzorak		**2. uzorak**		**3. uzorak**	
Ocjena	**Rang**	**Ocjena**	**Rang**	**Ocjena**	**Rang**
3	1	11	12	6	4.5
6	4.5	8	8.5	9	10
4	2	10	11	5	3
7	6.5	7	6.5	8	8.5
$\bar{R}_i$	**3.5**		**9.5**		**6.5**

R

```
> podacianova
   ocjena metoda
1       3     m1
2       6     m1
3       4     m1
4       7     m1
5      11     m2
6       8     m2
7      10     m2
8       7     m2
9       6     m3
10      9     m3
11      5     m3
12      8     m3
> kruskal.test(ocjena ~ metoda , data = podacianova )

        Kruskal-Wallis rank sum test

data:  ocjena by metoda
Kruskal-Wallis chi-squared = 5.5972, df = 2, p-value = 0.0609

>
```

'Post-hoc' test.

```
> pairwise.wilcox.test(podacianova$ocjena,
podacianova$metoda)
```

Test predznaka

- Dizajn **prije i poslije**
- Dva sparena uzorka.
- X - obilježje prije tretmana

  Y - obilježje poslije tretmana
- Uzorci:

  $$X_1, X_2, \ldots, X_n$$

  $$Y_1, Y_2, \ldots, Y_n$$
- Promatra se razlika: $\quad D_i = Y_i - X_i$
- Ako tretman pozitivno utječe

  $\longrightarrow$ očekuje se veći broj pozitivnih vrijednosti D_i.
- Ako tretman negativno utječe

  $\longrightarrow$ očekuje se veći broj negativnih vrijednosti D_i.

- S^+ - broj pozitivnih razlika

  S^- - broj negativnih razlika
- Ukoliko nema utjecaja tretmana S^+ i S^- bi trebala biti podjednaki.
- Ako je distribucija obilježja prije i poslije tretmana ista, statistika S^+ distribuirana je prema binomnoj $B(n, 0.5)$ razdiobi.
- Za veliki uzorak može se koristiti normalna aproksimacija:

$$z = \frac{S^+ - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \quad \sim N(0,1)$$

- Ukoliko se pojedine vrijednosti prije i poslije tretmana poklapaju ($X_i = Y_i$, tj. $D_i = 0$), te opservacije treba isključiti iz uzorka.

Primjer. U istraživanju utjecaja stimulansa na krvni tlak, istraživači su dvanaestorici pacijenata dali stimulans. Svakom pacijentu je krvni tplak izmjeren prije i poslije davanja stimulansa.

Postoji li opravdanje za tvrdnju da stimulans povećava krvni tlak?

Pacijent	Krvni tlak Prije (X)	Krvni tlak Poslije (Y)
1	120	128
2	124	131
3	130	131
4	118	127
5	140	132
6	128	125

Pacijent	Krvni tlak Prije (X)	Krvni tlak Poslije (Y)
7	140	141
8	135	137
9	126	118
10	130	132
11	126	129
12	127	135

Krvni tlak		
Prije (X)	Poslije (Y)	Razlika (D)
120	128	8
124	131	7
130	131	1
118	127	9
140	132	-8
128	125	-3

Krvni tlak		
Prije (X)	Poslije (Y)	Razlika (D)
140	141	1
135	137	2
126	118	-8
130	132	2
126	129	3
127	135	8

$$S^+ = 9 \quad S^- = 3$$

R.

R Console

```
> prije <- c(120,124,130,118,140,128,140,135,126,130,126,127)
> poslije <- c(128,131,131,127,132,125,141,137,118,132,129,135)
>
> d=poslije-prije
> d
 [1]  8  7  1  9 -8 -3  1  2 -8  2  3  8
>
> binom.test(sum(d > 0), length(d), p=0.5, alternative="two.sided")

        Exact binomial test

data:  sum(d > 0) and length(d)
number of successes = 9, number of trials = 12, p-value =
0.146
alternative hypothesis: true probability of success is not equal to 0$
95 percent confidence interval:
 0.4281415 0.9451394
sample estimates:
probability of success
                  0.75

>
```

Wilcoxonov test rangova s predznakom

- Dizajn **prije i poslije**
- Slično kao i test predznaka jedino se u obzir uzimaju rangovi razlika $D_i = Y_i - X_i$.
- Razlike D_i se rangiraju po apsolutnim vrijednostima (bez obzira na predznak)
- Ukoliko nema razlike između obilježja (populacija) X i Y, tada bi suma rangova za negativne razlike trebala biti podjednaka kao i suma rangova za negativne razlike.
- **Statistika**

  $T^+ =$ zbroj rangova pozitivnih razlika

- Za veliki uzorak distribucija od T^+ je približno normalna s

$$\begin{aligned} E(T^+) &= \frac{n(n+1)}{4} \\ Var(T^+) &= \frac{n(n+1)(2n+1)}{24} \end{aligned}$$

- Statistika:

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

- **Korekcija za izjednačene rangove:**

$$C = \sum_{j=1}^{g} \left(t_j^3 - t_j \right).$$

  - g - broj različitih vrijednosti za koje su rangovi izjednačeni
  - t_j - broj opservacija s istom j-tom vrijednošću ($j = 1, \ldots, g$)

- Korigirana statistika:

$$z = \frac{T^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{C}{48}}}$$

Primjer s krvnim tlakom.

Krvni tlak		
Prije (X)	**Poslije** (Y)	**Razlika** (D)
120	128	8
124	131	7
130	131	1
118	127	9
140	132	-8
128	125	-3

Krvni tlak		
Prije (X)	**Poslije** (Y)	**Razlika** (D)
140	141	1
135	137	2
126	118	-8
130	132	2
126	129	3
127	135	8

Određujemo rangove:

1	1	2	2	3	3	7	8	8	8	8	9
1.5	1.5	3.5	3.5	5.5	5.5	7	9.5	9.5	9.5	9.5	12

Krvni tlak			
Prije (X)	**Poslije** (Y)	**Razlika** (D)	**Rang**
120	128	8	9.5
124	131	7	7
130	131	1	1.5
118	127	9	12
140	132	-8	9.5
128	125	-3	5.5

Krvni tlak			
Prije (X)	**Poslije** (Y)	**Razlika** (D)	**Rang**
140	141	1	1.5
135	137	2	3.5
126	118	-8	9.5
130	132	2	3.5
126	129	3	5.5
127	135	8	9.5

$$T^{+} = 9.5 + 7 + 1.5 + 12 + 1.5 + 3.5 + 3.5 + 5.5 + 9.5 = 53.5$$

R.

R Console

```
> wilcox.test(prije, poslije, paired = TRUE)

        Wilcoxon signed rank test with continuity correction

data:  prije and poslije
V = 24.5, p-value = 0.2697
alternative hypothesis: true location shift is not equal to 0

Warning message:
In wilcox.test.default(prije, poslije, paired = TRUE) :
  cannot compute exact p-value with ties
> |
```

Napomena. $V(=T)$ je različit jer prikazuje zbroj rangova negativnih razlika $(24.5 = 12 \cdot 13/2 - 53.5$)

Spearmanov koeficijent korelacije

- Koeficijent korelacije mjeri povezanost dvije varijable.
- Pearsonov koeficijent korelacije mjeri **linearnu** povezanost dvije varijable.
- Spearmanov koeficijent korelacije mjeri koreliranost rangova.
- Alternativni naziv: **Spearmanov koeficijent korelacije rangova**.

- Promatramo dva obilježja u populaciji: X i Y.
- Uzorak veličine n:

$X: \quad X_1, X_2, \ldots, X_n$

$Y: \quad Y_1, Y_2, \ldots, Y_n$

- Za svaki uzorak odredimo rangove (unutar pojedinog uzorka):

$R: \quad R_1, R_2, \ldots, R_n$

$P: \quad P_1, P_2, \ldots, P_n$

- Spearmanov koeficijent korelacije

$$r_S = \frac{1}{n-1} \frac{\sum_i \left(R_i - \bar{R}\right)\left(P_i - \bar{P}\right)}{S_R \cdot S_P}$$

- Jer su rangovi vrijednosti od 1 do n:

$$\bar{R} = \bar{P} = \frac{1}{n}(1 + 2 + \ldots + n) = \frac{n+1}{2}$$

i

$$S_R^2 = S_R^2 = \sum_i (R_i - \bar{R}_i)^2 = \sum_i \left(i - \frac{n+1}{2}\right)^2 = \frac{n(n^2-1)}{12}.$$

- Spearmanov koeficijent korelacije

$$r_S = \frac{1}{n-1} \frac{\sum_i \left(R_i - \frac{n+1}{2}\right)\left(P_i - \frac{n+1}{2}\right)}{\frac{n(n^2-1)}{12}}$$

- Ako s D_i označimo razliku rangova: $D_i = R_i - P_i$, Spearmanov koeficijent korelacije je

$$r_S = 1 - \frac{6 \cdot \sum_i D_i^2}{n \cdot (n^2 - 1)}$$

- $-1 \leq r_S \leq 1$
- Interpretacija koeficijenta je analogna interpretaciji Pearsonovog koeficijenta, uz iznimku da se ne radi o linearnoj zavisnosti.
- Vrijednosti Spearmanovog koeficijenta r_S blizu 1 znači da su veće vrijednosti obilježja X pridružene većim vrijednostima obilježja Y.
- Za r_S blizu -1, većim vrijednostima obilježja X pridružene su manje vrijednosti obilježja Y.

Testiranje hipoteze o korelaciji rangova.

- ρ_S - Spearmanov koeficijent korelacije za populaciju.
- Može se testirati hipoteza o nepostojanju korelacije između rangova u populaciji:

$$H_0 : \quad \rho_S = 0.$$

- Za testiranje hipoteze kao statistika se koristi Spearmanov koeficijent korelacije r_S.

Primjer. (iz Pearsonovog koeficijenta korelacije) Na osnovu uzorka od 10 osoba procijenite koeficijent korelacije za visinu i težinu.

Visina (cm)	**Težina (kg)**
183	76
163	52
180	61
168	64
160	52
157	48
185	94
155	46
193	118
173	57

Računanje rangova.

Visina (cm)	Težina (kg)	Rang Visina	Rang Težina
183	76	8	8
163	52	4	3.5
180	61	7	6
168	64	5	7
160	52	3	3.5
157	48	2	2
185	94	9	9
155	46	1	1
193	118	10	10
173	57	6	5

Sada na rangovima izračunamo Pearsonov koeficijent korelacije.

R

Izbor testa

R Console

```
> visina <- c(183,163,180,168,160,157,185,155,193,173)
> tezina <- c(76,52,61,64,52,48,94,46,118,57)
> cor.test(visina,tezina, method="spearman")

        Spearman's rank correlation rho

data:  visina and tezina
S = 6.519, p-value = 1.016e-05
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.9604908

Warning message:
In cor.test.default(visina, tezina, method = "spearman") :
  Cannot compute exact p-value with ties
> 
```

KRAJ PREDAVANJA

H V A L A!