

Tau (τ) de Kendall

Soient deux échantillons de taille n décrits par deux variables X et Y (scores ou classements, à « l'individu » i correspond le couple (x_i, y_i)).

Ces n couples sont classés par ordre croissant par rapport à la variable X .

Le coefficient τ , appelé coefficient de corrélation des rangs de Kendall et noté r_k est calculé en tenant compte des différences de classement des deux séries.

Tau (τ) de Kendall

x	y_1	y_2
1	1	1
2	2	3
3	3	2

$n_a=3$

$n_d=0$

$n_a=2$

$n_d=1$

$$n_a + n_d = \text{nombre de paires} : \frac{n(n-1)}{2}$$

$$r_k = \frac{n_a - n_d}{\frac{n(n-1)}{2}} \in [-1; 1]$$

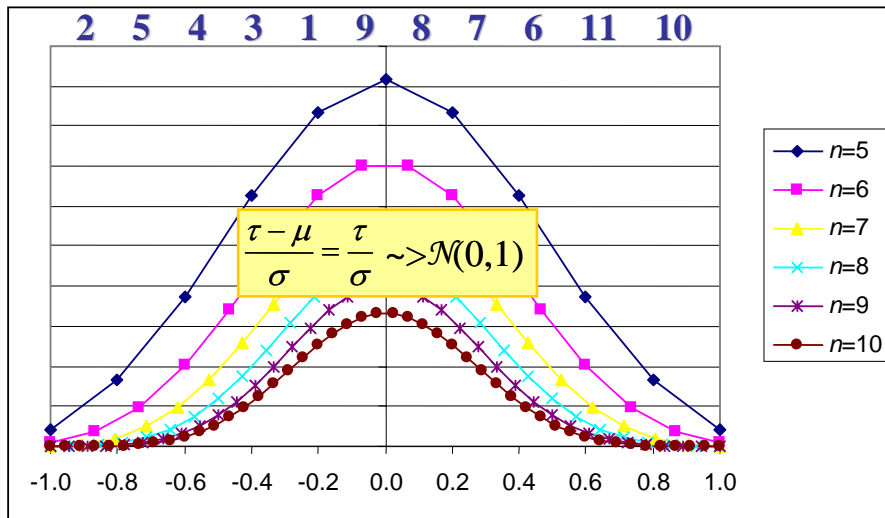
Tau (τ) de Kendall

Soient deux séries, y_1 et y_2 :

		$n(n-1)/2$			r_k		
		6	10	15	4	5	6
y_1	0	62.50%	59.17%		0.00	0.00	
	1			50.00%			0.07
	2	37.50%	40.83%		0.33	0.20	
	3			35.97%			0.20
	4	16.67%	24.17%		0.67	0.40	
	5			23.47%			0.33
	6	4.17%	11.67%		1.00	0.60	
	7			13.61%			0.47
	8		4.17%			0.80	
	9			6.81%			0.60
	10		0.83%			1.00	
	11			2.78%			0.73
	12						0.87
	13			0.83%			
	14						1.00
15			0.14%				

		4	5	6
y_2	1			
	4			
	2			
	3			
	5			
	6			

Tau (τ) de Kendall



Si $n > 10$ alors τ peut être approximé par une loi normale de paramètres : $\mu = 0$ et $\sigma = \sqrt{\frac{2(2n+5)}{9n(n-1)}}$