The Kendall's τ

计算方法

- Compute rank for X and Y 计算 Rank
- Classify each point 对每个点的分类
  - Positive: Rx > Ry 正的
  - Negative: Rx < Ry 负的
  - Zero: Rx = Ry 零（最后不会考虑的）
- Sum by category 按照类别统计次数
  - p: number of positive points 正类的个数
  - n: number of negative points 负类的个数
  - z: number of zeros points 零类的个数
  - $N = p + n + z$
- 逻辑
  - Positive 类别里互相是 concordant
  - Negative 类别里互相是 Concordant
  - Positive 和 Negative 之间的是 Discordant
- Pairs 计算对数
  - Concordant: $C_n^2 + C_p^2$ 一致的对数
  - Discordant: $n \times p$ 不一致的对数
  - Total: C_N^2 总对数
- Metric
  - $\frac{C_n^2 + C_p^2 - n \times p}{C_N^2}$

Pair Combination

	Positive	Negative	Zero
Positive	C_p^2		
Negative	$n \times p$	C_n^2	
Zero	$z \times p$	$z \times n$	C_z^2

组合只看下三角：红色是 concordant pairs，蓝色是 discordant pairs，黑色是被忽略的

需要证明下三角之和是所有的对数

$C_p^2 + C_n^2 + n \times p + z \times p + z \times n + C_z^2 = C_N^2$

$=> C_p^2 + C_n^2 + n \times p + z \times (N - z) + C_z^2 = C_N^2$

$=> p \times (p - 1) + n \times (n - 1) + 2 \times p \times n + 2 \times z \times (N - z) + z \times (z - 1)$
$= N \times (N - 1)$

$=> p^2 - p + n^2 - n + 2 \times p \times n + 2 \times z \times (N - z) + z^2 - z = N^2 - N$

$=> p^2 + 2 \times p \times n + n^2 + 2 \times z \times (N - z) + z^2 - n - p - z = N^2 - N$

$=> (p + n)^2 + 2 \times z \times (N - z) + z^2 = N^2$

$=> (N - z)^2 + 2 \times z \times (N - z) + z^2 = N^2$

$=> (N - z + z)^2 = N^2$

A risk manager gathers five years of historical returns to calculate the Kendall τ correlation coefficient for stocks *X* and *Y*. The stock returns for *X* and *Y* from 2010 to 2014 are as follows:

Year	*X*	*Y*
2010	5.0%	−10.0%
2011	50.0%	−5.0%
2012	−10.0%	20.0%
2013	−20.0%	40.0%
2014	30.0%	15.0%

X	Y	Rx	Ry
-20%	40	1	5
-10%	20	2	4
5%	-10	3	1
30%	15	4	3
50%	-5	5	2

What is the Kendall τ correlation coefficient for the stock returns of *X* and *Y*?

A. −0.3.

B. −0.2.

C. 0.4.

D. 0.7.

nc = 2   nd = 8

$\tau = \frac{2-8}{5\times4/2} = \frac{-6}{5\times2} = \cancel{-0.4}\ -0$

解答

- **计算类别**

Year	X	Y	类别 (X>Y?)	Rank X	Rank Y	Rank X> Rank Y
2010	5%	-10%	1	3	1	-1
2011	50%	-5%	1	5	2	1
2012	-10%	20%	-1	2	4	-1
2013	-20%	40%	-1	1	5	-1
2014	30%	15%	1	4	3	1

- 统计类别
  - Positive (1)：2 个
  - Negative (-1)：3 个
- Pair 数
  - Coordinate: $C_n^2 + C_p^2 = C_2^2 + C_3^2 = 1 + 3 = 4$
  - Discordant: $n \times p = 3 \times 2 = 6$
  - Total: $C_N^2 = C_5^2 = 10$
- Metric
  - $\frac{C_n^2 + C_p^2 - n\times p}{C_N^2} = \frac{4-6}{10} = -0.2$

Calculate the Kendall τ correlation coefficient for the stock returns of *X* and *Y* listed in Figure 3.

Figure 3: Ranked Returns for Stocks *X* and *Y*

Year	*X*	*Y*	*X Rank*	*Y Rank*
2012	–20.0%	10.0%	1	2
2014	–10.0%	30.0%	2	4
2010	25.0%	–20.0%	3	1
2013	40.0%	20.0%	4	3
2011	60.0%	40.0%	5	5

Answer:

Begin by comparing the rankings of *X* and *Y* stock returns in columns four and five of Figure 3. There are five pairs of observations, so there will be ten combinations. Figure 4 summarizes the pairs of rankings based on the stock returns for *X* and *Y*. There are two concordant pairs, four discordant pairs, and four pairs that are neither concordant nor discordant.

Figure 4: Categorizing Pairs of Stock *X* and *Y* Returns

Concordant Pairs	*Discordant Pairs*	*Neither*
{(1,2),(2,4)}	{(1,2),(3,1)}	{(1,2),(5,5)}
{(3,1),(4,3)}	{(1,2),(4,3)}	{(2,4),(5,5)}
	{(2,4),(3,1)}	{(3,1),(5,5)}
	{(2,4),(4,3)}	{(4,3),(5,5)}

Kendall's τ can then be determined as –0.2:

$$\tau = \frac{n_c - n_d}{n(n-1)/2} = \frac{2-4}{5(5-1)/2} = -0.2$$

- **计算类别**

Year	X	Y	类别 (X>Y?)	Rank X	Rank Y	Rank X> Rank Y
2012	-20%	10%	-1	1	2	-1
2014	-10%	30%	-1	2	4	-1
2010	25%	-20%	1	3	1	1
2013	40%	20%	1	4	3	1
2011	60%	40%	1	5	5	0

- 统计类别

- Positive (1)： 2 个
- Negative (-1)： 2 个
- Zero(0): 1 个

- Pair 数
  - Coordinate: $C_n^2 + C_p^2 = C_2^2 + C_2^2 = 1 + 1 = 2$
  - Discordant: $n \times p = 2 \times 2 = 4$
  - Total: $C_N^2 = C_5^2 = 10$
- Metric
  - $\frac{C_n^2 + C_p^2 - n \times p}{C_N^2} = \frac{2-4}{10} = -0.2$