Quantifying dis-/concordance between ranked lists

Cape vs SWA publication

Ruan van Mazijk

2018-12-21

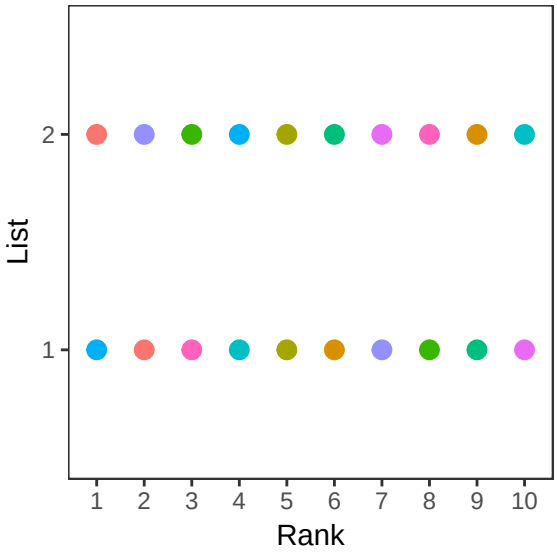
Introduction

Suppose we have a dataset with two different rankings of the same set of categories, e.g.:

Category	Ranking no. 1	Ranking no. 2
a	2	1
b	6	9
c	5	5
d	8	3
e	9	6
f	4	10
g	1	4
h	7	2
i	10	7
j	3	8

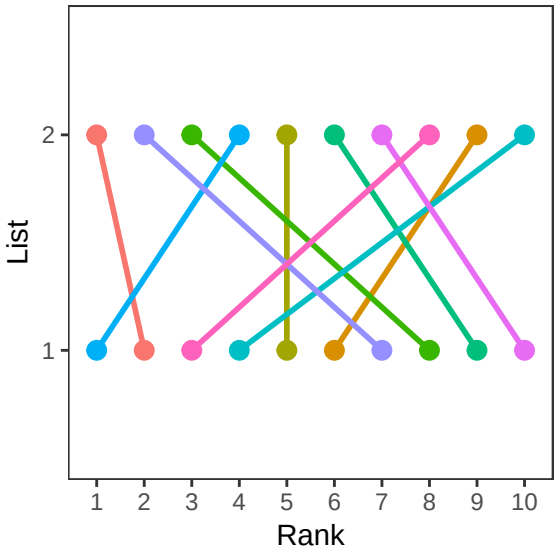
Visualising rank change

How do we visualise the differences in ranking between these two lists? Well, let us visualise the data on their own first:



Here, each point is an observation, and points have been organised by rank and coloured by their category.

Now, let us connect points belonging to the same category with lines:



Quantifying rank change

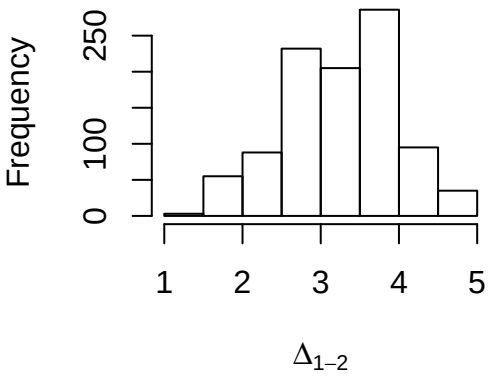
How do we quantify the change in these categories' ranks moving from list no. 1 to 2? Generally, we wish to describe how many positions a category has changed in rank when moving from one list to the other. This can be represented mathematically rather intuitively. Formally, let δ_i be difference between the i^{th} category's position in one list (p_1) versus the other (p_2). Here, we do not wish to distinguish between positive and negative changes in position, so we shall take the absolute value of all δ_i. Thus, the average change in rank between lists no. 1 and 2 (Δ_{1-2}) is as follows:

$$\Delta_{1-2} = \overline{|\delta|} = \sum_{i=1}^{n} |\delta_i| = \sum_{i=1}^{n} |p_{1i} - p_{2i}|$$

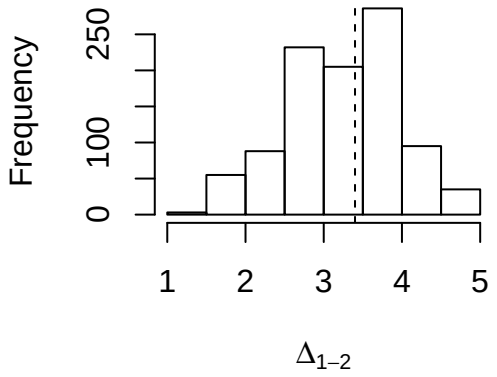
In the example data above, $\Delta_{1-2} = 3.4$.

Drawing inference about rank change

How do we know if some value of Δ_{1-2} differs from what chance would produce? Let us compute a **randomised null** average change in rank between the two lists by randomly re-ranking the categories in both lists 999 times:



We can rank our observed Δ_{1-2} amongst these 999 permutations:



We can derived from this a P-value as follows (appropriately, using **ranks**):

```
ranked_Deltas <- rank(c(obs, null))
obs_Delta_rank <- ranked_Deltas[1]
p_value <- obs_Delta_rank / (999 + 1)
p_value
```

```
## [1] 0.535
```