Shimane Journal of Policy Studies
Vol. 4 (February 2003)

Comparison of Life Styles by Testing the Similarity of Pairs of Graphs※

FUJIWARA Masago

Activity rates in a single day have been relatively neglected data in time use studies. I have been focusing mainly on developing methods for analyzing them. In this paper, I try to present a new idea to examine the way of discerning the similarity or dissimilarity between a pair of line graphs that chart activity rates of every time slot in a day.

We are analyzing them, focusing on the difference between both of their shape of wave motion and their height of wave. To accomplish this task, we introduce Chi-square test and F-test into our analysis.

1. Analyzing the shape of wave motion

In most time use studies, we draw a step chart to illustrate a series of activity rates data of a day. But, I use instead a line graph throughout this paper, as is shown in Figure 1. Furthermore, I make use of the length of 20 activity rates data for simple and plain explanation, although the length of activity rates data of time use research is generally 48 or 96.

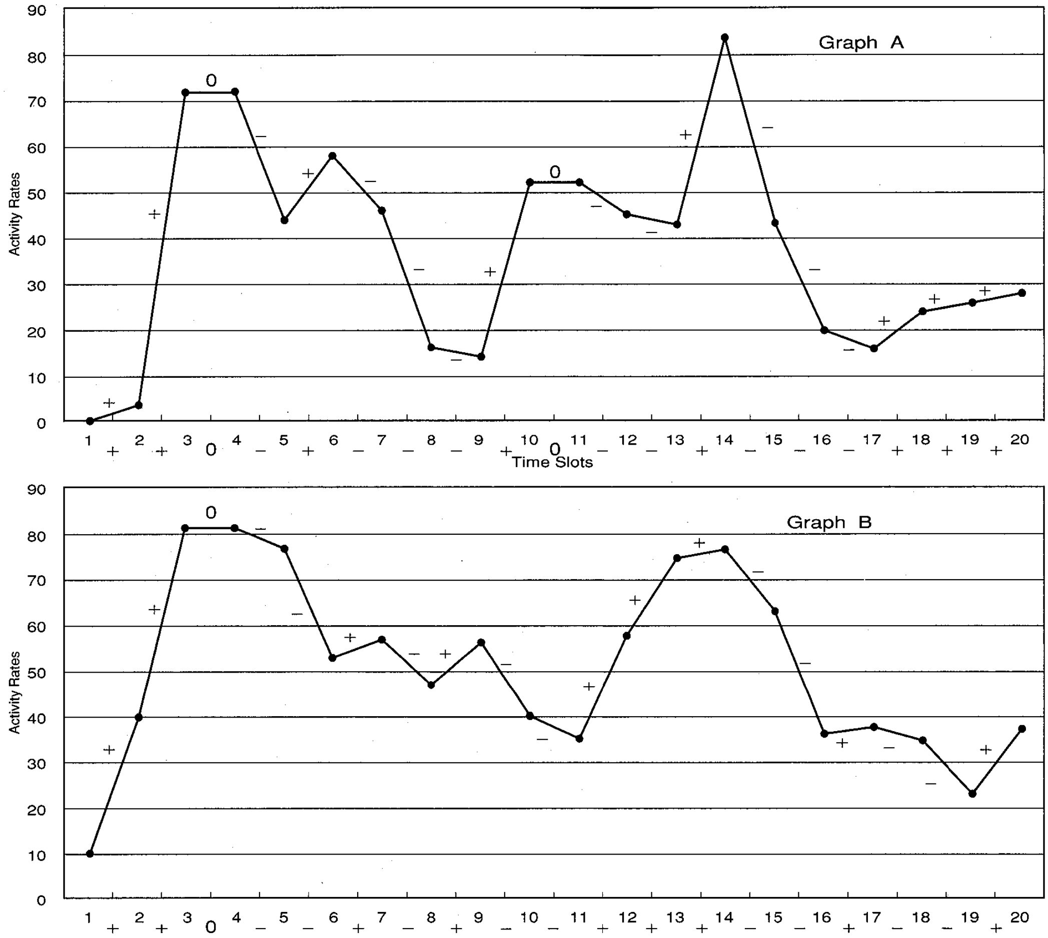
Figure 1 is drawn for analyzing a similarity of a pair of graphs. When one of the slopes of a line graph is ascending, we put a positive sign on it. On the contrary, if another slope is descending, a negative sign is put on it. And, in a case of a flat slope, a sign of 0 is given to it.

A basic idea of analysis[1)] is that, as long as a sign of Graph A at a point

---

※ This paper was presented at the International Association for Time Use Research (IATUR) Conference 2002, 15-18 October, held at the Technical University of Lisbon, Portugal.

Figure 1 How to measure the similarity of wave motion between two line graphs?



Notes: 1. When a slope is rising, it gets a positive sign.
2. When a slope is declining, it obtains a negative sign.
3. When a slope is flat, it has a sign of 0.

Table 1 List of signs between Graph A and Graph B.

Slope	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Graph A	+	+	0	−	+	−	−	−	+	0	−	−	+	−	−	−	+	+	+
Graph B	+	+	0	−	−	+	−	+	−	−	+	+	+	−	−	+	−	−	+
Product	+	+	+	+	−	−	+	−	−	−	−	−	+	+	+	−	−	−	+

Notes: Products of the third row are following the rule listed below.
1. If two signs are positive or negative, their product is positive.
(1) + and + → + (2) - and - → +
2. If one is positive and the other is negative, their product is negative.
+ and - → -
3. If two signs are zero, their product is positive, because they can be regarded as showing the same move.
0 and 0 → +
4. If one is zero and the other is positive or negative, their product is negative, because they have a different move.
(1) 0 and + → - (2) 0 and - → -

Table 2 Summary of products of Table 1

Category of sign	+ (Positive sign)	− (Negative sign)
Frequency	9	10

shows the same as that of Graph B, we regard it as a similar move. On the other hand, when it has a different sign, we judge that it has a different move. Some logical rules of product between a sign of Graph A and that of Graph B are shown in the note of Table 1.

Let us suppose three extreme cases. If a pair of line graphs shows completely the same sign at every slope, their signs of product show a positive sign. On the other hand, if they show a different sign in every slope, all their products give a negative sign on every slope. And, when they have no relation at all, one half of products give a positive sign and another half of products show a negative sign.

Then, I propose the use of a Chi-square test to examine the relation of two graphs. Here, we set the following null hypothesis.

Null hypothesis: the wave motion of Graph A and that of Graph B have no relation at all.

As is clear, the most general formula for Chi-square is following.

$$\chi^2 = \sum_{i=1}^{n} \frac{(O-E)^n}{E}$$

In the formula, n is the number of cells in the table, O is the observed frequency in each cell, and E is the expected frequency in each cell.

In our case, we have two cells, namely, a cell of positive sign and a cell of negative sign, shown in Table 2. The degree of freedom is 1.

$$\chi^2 = \frac{(\text{frequency of positive sign} - E)^2}{E} + \frac{(\text{frequency of negative sign} - E)^2}{E}$$

Following our null hypothesis, the frequency of positive sign and negative sign are expected to show the same frequency. The expected frequency is

9.5, because it is a half of the total frequency 19. As is clear in the Table 2, the frequency of positive sign is 9, and that of negative frequency is 10. Accordingly, the process and the result of calculation are as follows.

$$\chi^2 = \frac{(9-9.5)^2}{9.5} + \frac{(10-9.5)^2}{9.5} = 0.05$$

$$\phi = 1$$

The critical value of 5% is 3.84 in case of one degree of freedom. If our Chi-square value is at or beyond 3.84, our null hypothesis can be rejected. Its value shows 0.05 and is below 3.84. Therefore, we can conclude that our null hypothesis can be verified and supported.

2. Analyzing the height of wave

Let us take another example. The Figure 2 shows another pair of graphs. Graph A is the same as the Graph A of Figure 1. Graph A1 is charted by one half the activity rates of Graph A. The wave motion of Graph A is the same as that of Graph A1.

As is clear in Figure 2 a pair of signs at every slope has a same sign at every slope. Accordingly, their product is always positive (See the bottom of

Figure 2 A pair of Graphs with the same signs and different variance

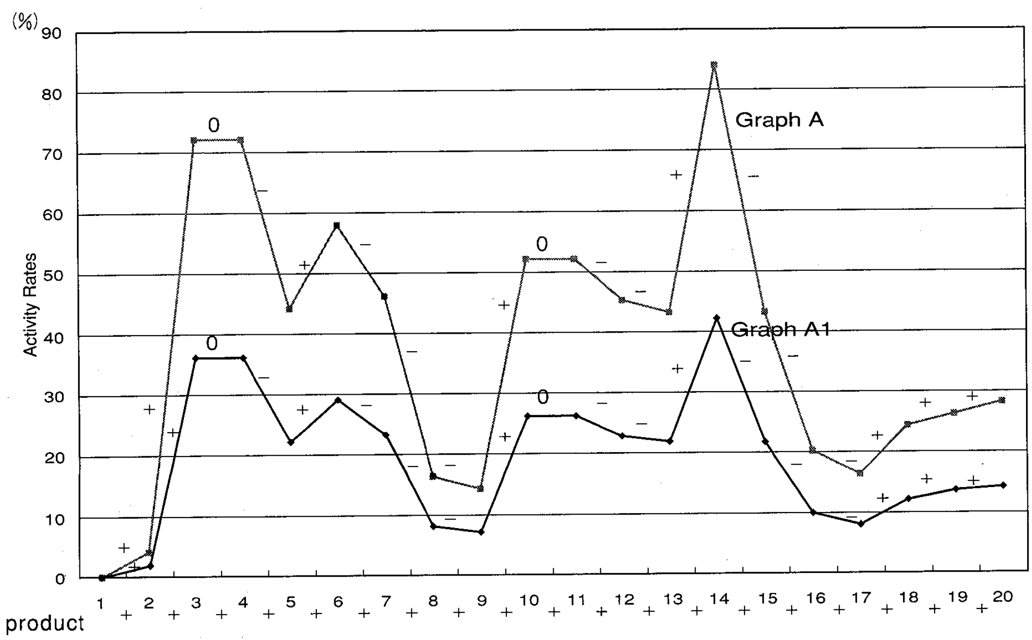


Table 3 Summary of products in Figure 2

Category of sign	+ (Positive sign)	− (Negative sign)
Frequency	19	0

Figure 2, and Table 3). Chi-square value is calculated in the following formula.

$$\chi^2 = \frac{(19-9.5)^2}{9.5} + \frac{(0-19.5)^2}{9.5} = 19.00$$

$$\phi = 1$$

In this case also, we set the following null hypothesis as before.

Null hypothesis: the wave motion of Graph A and that of Graph A1 have no relation at all.

Chi-square value shows the maximum value of 19.00 that is far beyond 3.84 in case of critical value of 5 percent. We can conclude that our null hypothesis cannot be verified.

Let us think for a while. As for a shape of wave, Graph A and Graph A1 show the biggest similarity. But, when we turn to the height of wave, namely, the value of activity rates of Graph A and Graph A1 at every time slot, it is quite clear that there are great difference between them.

Through this case, it is apparent that we should also pay attention to the variance of activity rates for examining the similarity of a pair of line graph.

Here, we introduce F-test into our analysis. At first, we divide one of two variance estimates by the other to obtain a F-value. If the F-value surpasses the critical value in the F-table, we can conclude that their variance estimates differ significantly.

The variance estimate of activity rates of Graph A ($=V_1$) is 543.52, while that of Graph A1 ($=V_2$) is 135.58. Here, let us calculate the F-value (F-ratio) with V_1 as numerator, and V_2 as denominator. The length of vector data of Graph A and that of Graph B is respectively 20. Accordingly, both have a degree of freedom of 19.

We set the following hypothesis.

Hypothesis: Sample V_1 and V_2 are extracted from the same population.

$$V_1 = 543.52 \quad V_2 = 135.58$$

$$F = \frac{\frac{V_1}{\sigma_1^2}}{\frac{V_2}{\sigma_2^2}}$$

As we take as the hypothesis that the population of V_1 and that of V_2 are the same, we can set as follows.

$$\sigma_1 = \sigma_2$$

Therefore,

$$F = \frac{V_1}{V_2} = \frac{543.52}{135.58} = 4.01$$

As the degree of freedom is 19 in both numerator and denominator, the critical value of 2.5 percent under two-tailed test is as follows.

$$F_{19}^{19}(0.025) = 2.53$$

$$F_{19}^{19}(0.975) = 0.40$$

If our F-value (F-ratio) is within the following range, our hypothesis can be supported.

$$F_{19}^{19}(0.975) \leq F \leq F_{19}^{19}(0.025)$$

$$0.40 \leq F \leq 2.53$$

As our F-value shows 4.01, it is beyond the range. It goes without saying that the reciprocals, 0.25 (=1/4.01), are also outside and below the range.

Therefore, we can judge that our hypothesis cannot be verified.

3. Synthesizing analytical methods

(1) Frame of reference for statistical verification

So far, we have tried to introduce Chi square test and F-test into our analysis.

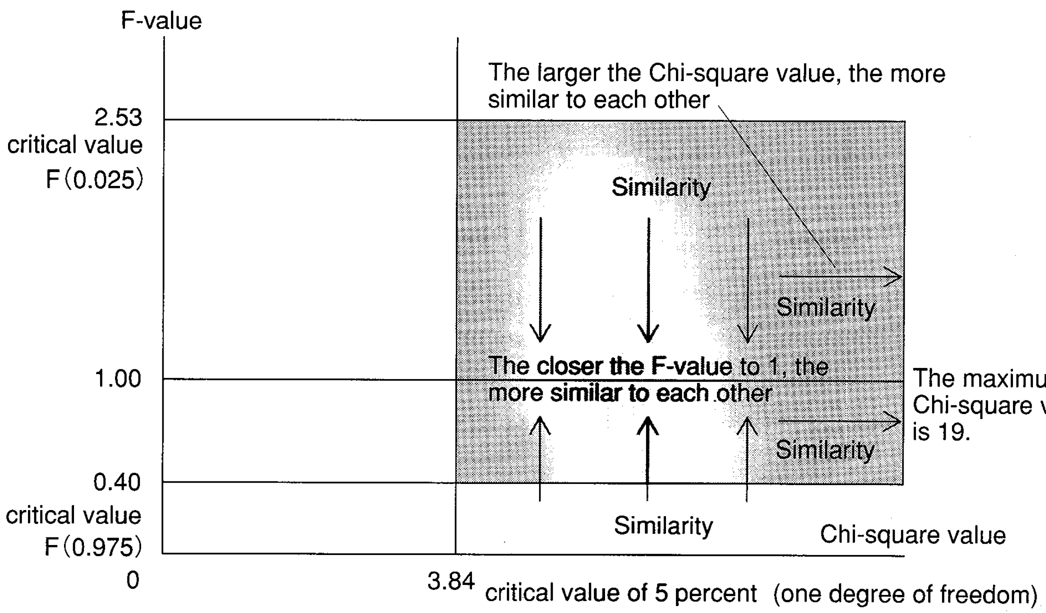
As for the relation between Graph A and Graph A1, through Chi-square test, we could deny our null hypothesis, and through F-test, we concluded that our samples are not extracted from the same population. In sum, although Chi-square test verifies the similarity of wave motion of two graphs, F-test denies their similarity. Therefore, we may conclude that there exists no similarity between them.

So far as both Chi-test and F-test verify the similarity of two graphs, we can say that they are similar to each other. Let us visualize our logical structure. Figure 3 makes clear our Frame of reference for statistical verification. It serves us to identify the similarity of a pair of graphs.

The range, shadowed in gray, is a zone where similarity of a pair of graphs can be verified by both Chi-square test and F-test. When we compare one with another line graph, as far as both the Chi-square value and the F-value are within the range, in other words, as long as a pair of values satisfies two statistical conditions at the same time, we can verify that there exists an ob-

Figure 3 Range of similarity verified by Chi-square test and F-value test between a pair of activity rates vector of length 20

(The degree of freedom of F-value is 19 in both numerator and denominator)



vious similarity between them.

(2) Introducing an idea of dissimilarity to our analytical concept

We can look into a value of Chi-square in every combination of positive and negative signs of products in Table 4.

So far, we have been discussing the similarity of two graphs. But, Table 4 makes us notice that we have also to pay attention to the dissimilarity between them. Furthermore, we can find the range of "no relationship" exists between similarity and dissimilarity.

Table 4 The combination of products, its Chi-square values, and degree of similarity
(in case of a pair of graphs, having 20 activity rates)

	⇦ strong relationship									weak relationship ⇨
	Range of similarity of a pair of graph (positive relation)						No relationship			
+	19	18	17	16	15	14	13	12	11	10
−	0	1	2	3	4	5	6	7	8	9
χ^2	19.00	15.21	11.84	8.89	6.37	4.26	2.58	1.32	0.47	0.05
+	0	1	2	3	4	5	6	7	8	9
−	19	18	17	16	15	14	13	12	11	10
χ^2	19.00	15.21	11.84	8.89	6.37	4.26	2.58	1.32	0.47	0.05
	Range of dissimilarity of a pair of graphs (negative relation)						No relationship			

Note: The critical value of 5 percent is 3.84 under one degree of freedom

Therefore, we have to revise our frame of reference for statistical verification from Figure 3 to Figure 4. Through looking in the Table 4, we notice the relation that the larger the Chi-square value, the more similar or dissimilar a pair of graph is.

In closing our methodological discussions, let us synthesize and arrange our findings, referring to Table 4.

(1) The larger the Chi-square value, the more similar or dissimilar a pair of graphs is, and the closer the F-value is to 1, the more similar a pair

of graphs is.

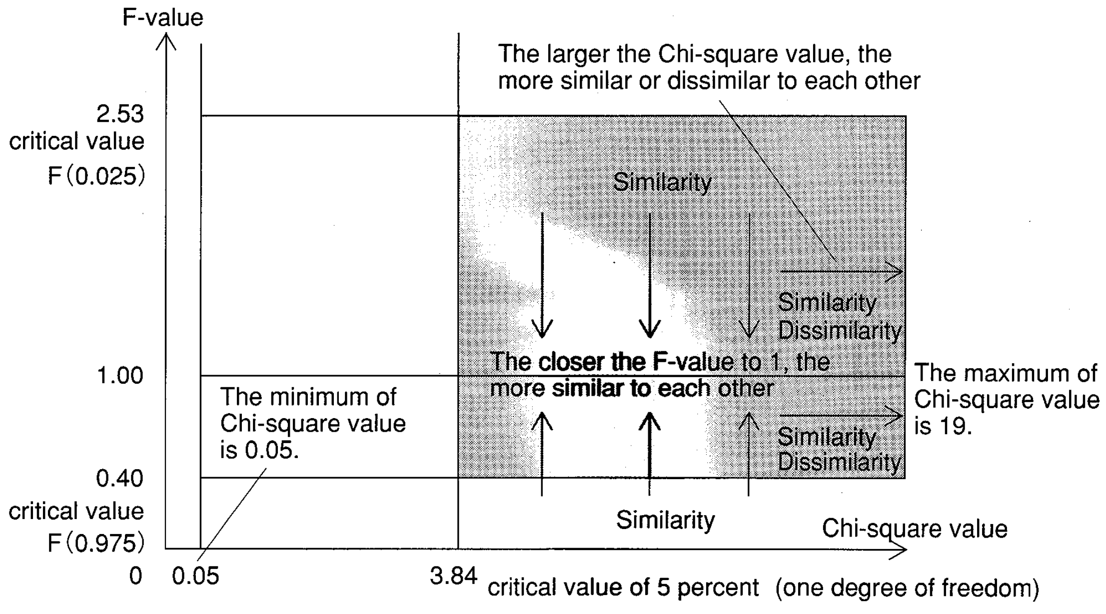
(2) We say there can exist a similarity or dissimilarity between a pair of graphs, as far as both their F-value and Chi-square value exist within the range of critical value of both F-test and Chi-square test.

4. Implications for future research

The essential points of my argument have been expressed in the proceeding pages. In ordinary time use studies, we use activity rates data having the length of 48 or 96. We should have dealt with such cases to extend our analytical logic. The future direction of this study will be devoted to empirical research using such real time use data. I hope that the result of this survey will contribute to the development of time use studies, although it might be just a small step.

Figure 4 Range of similarity and dissimilarity verified by Chi-square test and F-value test between a pair of activity rates vector of length 20

(The degree of freedom of F-value is 19 in both numerator and denomirator)



Acknowledgements:

I would like to thank Sean White of the University of Shimane for his help in proofreading this paper.

Notes

1) Regarding a basic idea for analyzing the shape of wave motion, I owe very much to SATO Makoto (1968, pp.167-169).

References

SATO Makoto 1968 *Suikeigaku no Susume.* Bluebacks-116. Tôkyô: Kôdansha.

Key words : Time Use Study, Activity Rates, Similarity of wave motion, chi-square Test, F-value Test, Degree of Similarity

(FUJIWARA Masago)