# Analysis on Skeptical Analysis 

Serge Daigno
29-06-2017

## Summary

This paper follows the analysis of a CSI \& Skeptical Inquirer member on the Grandmasters and Sports studies. Each of the arguments raised by this analyst is reviewed and clarified by the author of these researches.

## Introduction

First of all, I would like to thank the CSI member who spent some personal time to evaluate and give his opinion on the relevance of the studies. However, I think it is important to clarify some elements that were not taken into account in his analysis.

This article takes up each of the arguments that underestimate the actual value of the data and proposes new tables that confirm the robustness of results.

## Some Clarifications

Huge sample sizes will inflate the most trivial differences to impressive statistical significance.

This is not the case in this study. As example, this table shows how Nbs numbers are scaled down in order to calculate $p$ value $=.0023$ in the most conservative manner.

| Conjunctions | Mercury | Nbs |
| :---: | :---: | :---: |
| Yes | 92 | 4,173 |
| No | 193 | 12,846 |
| Total | 285 | 17,019 |
| Scaled down Nbs | $(4,173 * 285) / 17,019=69.881015$ | $(12,846 * 285) / 17,019=215.11898$ |

Therefore, this argument cannot be invoked in the CSI analysis.

The observed monthly distribution of births varies between countries, between periods, and even between different ethnic groups in the same country.

These factors have no significance in this study. The control groups data are based on RAE facts caused by the introduction of a cut-off date into sport drafts:
http://www.slate.com/articles/sports/sports nut/2008/04/the boys of late summer.html
Therefore, at least in this sports study, this argument cannot be invoked in the CSI analysis.

The occurrence of conjunctions is non-uniform due to planetary retrograde...

This is why studies provide a breakdown per period into Section 3. Percentages displayed for most periods confirm the resilience of conjunctions regardless the apparent motion of planets.

Therefore, this argument cannot be invoked in the CSI analysis.

An associated pitfall that Gauquelin occasionally (and inadvertently) did not avoid was to calculate chi-squared values using a control group for the expected values instead of the theoretical values.

Each of many randomized control groups will compute slightly different percentages (see Grandmasters and Sports annexes) that should not be exceeded by theoretical values. Such theoretical control group would not attenuate conclusions of the researches.

Therefore, this argument cannot be invoked in the CSI analysis.

One standard way to overcome this problem, and also the problem of mistaking statistical significance for practical significance, is to look at the actual association between results in a $2 \times 2$ contingency table.

This third argument refers to the Phi coefficient ( $\phi$ ) formula, used to evaluate the practical significance of a statistical result. Following examples explain how such formula can distort results of a statistical study:

The first table shows a coefficient of zero ( $\phi=0.0$ ), having both vessels $A$ and $B$ water-filled at $50 \%$. The last table, on the other hand, displays a very high coefficient ( $\phi=0.8$ ). Not surprisingly, it is easy to see that connected vessels A and B are water-filled in inverse proportion.

| $\phi=0.0$ | A | B |
| :---: | ---: | ---: |
| Water | 5 | 5 |
| Air | 5 | 5 |
| Volume | 10 | 10 |


| $\phi=0.4$ | A | B |
| :---: | ---: | ---: |
| Water | 7 | 3 |
| Air | 3 | 7 |
| Volume | 10 | 10 |


| $\phi=0.8$ | A | B |
| :---: | ---: | ---: |
| Water | 9 | 1 |
| Air | 1 | 9 |
| Volume | 10 | 10 |

In my astrological studies, the $\phi$ coefficient will always be low, since it is impossible for the comparative data to evolve in inverse proportion.

As another example, in the first table, container A contains 90\% apples, while container B contains $10 \%$ apples ( $\phi=0.8$ ). Despite the huge difference in $\%$, it is possible to almost nullify the $\phi$ coefficient ( $\phi=0.03$ ) by simply inflating the values in column B!

| $\phi=0.8$ | A | B |
| :---: | ---: | ---: |
| Apples | 9 | 1 |
| Oranges | 1 | 9 |
| Total | 10 | 10 |


| $\phi=0.08$ | A | B |
| :---: | ---: | ---: |
| Apples | 9 | 1,000 |
| Oranges | 1 | 9,000 |
| Total | 10 | 10,000 |


| $\phi=0.03$ | A | B |
| :---: | ---: | ---: |
| Apples | 9 | 10,000 |
| Oranges | 1 | 90,000 |
| Total | 10 | 100,000 |

Thus, in the same way, rather than comparing only the astrological conjunctions identified, it is possible to dilute the $\phi$ coefficient by injecting the complete data files into calculations.

Indeed, according to astrology they could now be denied being a grandmaster or team player, while ordinary people who happened to have one could claim automatic status.

Here is the conclusion of the chess study:
It would be wrong to conclude that the $\sigma \underset{\uparrow}{9}$ is the primary condition to succeed in chess. It illustrates perhaps rather some dynamic that occurs itself in anyone who enjoys activities involving logic.

## Grandmasters Annex

The control group (Table 1) has 56,480 random dates for the period 1880-1990. Date distribution peaks from $11.29 \%$ to $5.66 \%$. In the second table, dates are shifted within 300 days.

## Control Group

| J | F | M | A | M | J | J | A | S | O | N | D |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $11.29 \%$ | $9.54 \%$ | $9.36 \%$ | $8.28 \%$ | $7.54 \%$ | $6.40 \%$ | $5.66 \%$ | $6.63 \%$ | $7.27 \%$ | $8.46 \%$ | $9.08 \%$ | $10.48 \%$ |

$\sigma$ ૪ $\left( \pm 2.0^{\circ}\right)$ with $\odot \odot$

| GM vs Shifted days | Data | ¢ | $\odot$ |
| :---: | :---: | :---: | :---: |
| Grandmasters | 190 | $\mathbf{4 8 . 4 2 \%}$ | $51.58 \%$ |
| +000 Shift | 5,443 | $37.31 \%$ | $62.69 \%$ |
| +060 Shift | 5,257 | $37.44 \%$ | $62.56 \%$ |
| +120 Shift | 5,207 | $38.24 \%$ | $61.76 \%$ |
| +180 Shift | 5,124 | $38.47 \%$ | $61.53 \%$ |
| +240 Shift | 5,336 | $37.69 \%$ | $62.31 \%$ |
| +300 Shift | 5,265 | $38.06 \%$ | $61.94 \%$ |

Percentages displayed for the shifted days does not differ significantly.

## Sports +6 Years Annex

The control group (Table 1) has 39,886 random dates for the period 1950-1990. Date distribution peaks from $11.58 \%$ to $5.83 \%$. In the second table, dates are shifted within 300 days.

## Control Group

| J | F | M | A | M | J | J | A | S | O | N | D |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $11.58 \%$ | $9.78 \%$ | $10.15 \%$ | $9.32 \%$ | $9.14 \%$ | $8.42 \%$ | $8.18 \%$ | $7.77 \%$ | $7.01 \%$ | $6.72 \%$ | $6.10 \%$ | $5.83 \%$ |



| Sports vs Shifted days | Data | § | $\odot$ | $\sigma^{\top}$ |
| :---: | :---: | :---: | :---: | :---: |
| Sports 1950-1990 | 726 | $34.99 \%$ | $30.58 \%$ | $\mathbf{3 4 . 4 4 \%}$ |
|  |  |  |  |  |
| +000 Shift | 3,782 | $38.82 \%$ | $33.69 \%$ | $27.50 \%$ |
| +060 Shift | 3,629 | $40.37 \%$ | $33.09 \%$ | $26.54 \%$ |
| +120 Shift | 3,656 | $39.88 \%$ | $32.85 \%$ | $27.27 \%$ |
| +180 Shift | 3,625 | $39.64 \%$ | $32.74 \%$ | $27.61 \%$ |
| +240 Shift | 3,615 | $40.66 \%$ | $33.36 \%$ | $25.98 \%$ |
| +300 Shift | 3,824 | $39.04 \%$ | $35.12 \%$ | $25.84 \%$ |

Percentages displayed for the shifted days does not differ significantly.

