Query by Admin: In the BBC’s The Sky at Night - Playing with a Clockwork Universe, of 15 Mar 2019, available at

https://www.youtube.com/watch?v=fUqylNKUnIk

Hannah Fry makes the following statement regarding the findings of M. Gauquelin:

“To mathematicians like me, the result is not so ground-breaking… he was looking through so many different correlations, thousands and thousands of birthdates, with ten planets and twelve positions in the sky, that he was bound to come across some spurious patterns. And that’s what the Mars effect is. He didn’t adjust for the number of combinations, and the result is just there by chance.”

Response by Mersenne: Probably the pundit was making a more sophisticated point that got edited out. There is certainly the issue of *multiple comparisons*, which I here reproduce from Wikipedia:

“There are 10 celestial bodies and 12 sectors for them to be in. Furthermore, there are 132 combinations of sector pairs and thus 1320 different combinations of a planet with two sectors. There is about a 25% chance to find at least one such combination (of one planet and two sectors) for a random dataset of the same size as Gauquelin’s that would yield a result with apparent statistical significance like the one obtained by Gauquelin.[11] This implies that after adjusting for multiple comparisons, the Mars effect is no longer statistically significant even at the modest significance level of 0.05 and is probably a false positive.”

I note that this example also brings attention to the numbers of signs and planets, so possibly this is the argument intended. But as things stand the statement is puzzling, so let’s consider both issues.

1. Rather than patterns *appearing* with large datasets, the larger the sample, the more the frequencies will tend towards the expected frequencies—that is, spurious patterns will disappear. If you throw a coin a few thousand times, yes, it is more common to find runs of ten heads—but there will be runs of ten tails as well, and the frequencies will increasingly tend to half heads, half tails.

2. Not all the data Gauquelin examined does give any kind of pattern. For example, on just as big a sample, Mercury's Aspects to the Ascendant don't

show any pattern for sportsmen. Other sceptics have used this as evidence *against* the effect.

3. The patterns that are found involve the specific planets that astrology expects- Mars for Sportsmen, Saturn for doctors, Venus for writers, and so on. The "spurious patterns" theory doesn't account for this. In fact, it also argues against the "multiple comparisons" argument, since patterns appearing as a result of chance wouldn't support any system. (It's true MG's work doesn't support the traditional Houses—but they are after all a bone of contention in the Astrological community too. But no Astrologer would disagree that Mars is the sportsman's planet! Observing just *some* effect is liable to be chance, but observing an expected Astrological outcome is a different ball-game.)

4. If the pattern is spurious, then it would not be shown in alternative datasets. This is a stronger argument against both objections, though our ability to test it is limited (reasons A-C below).

5. If the pattern is spurious, smaller subsets would show different patterns. Again a good argument against both.

I've returned to the set of sportsperson's horoscopes to investigate these last two points.

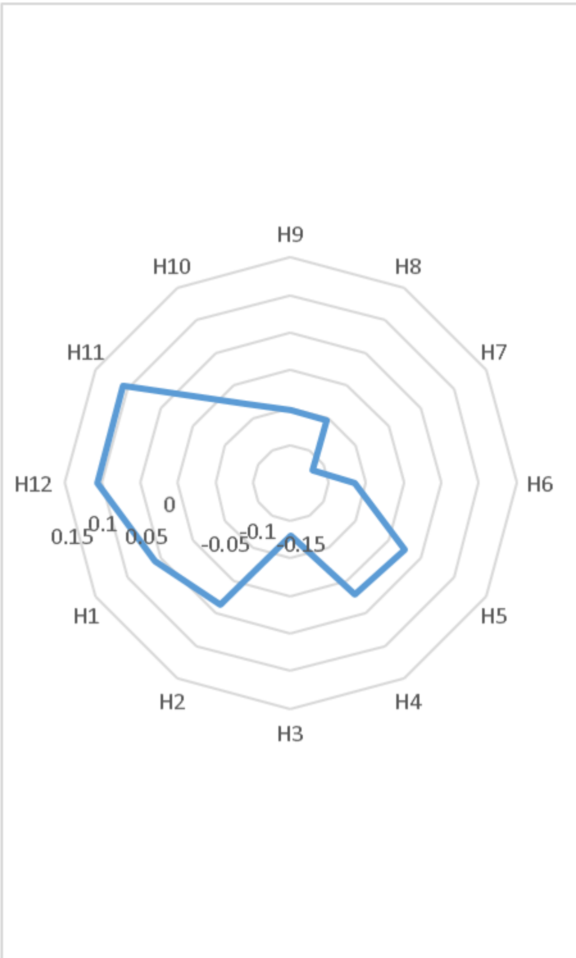
For point 4, here is the Gauquelin experiment repeated—rather crudely, it must be said, for the following reasons.

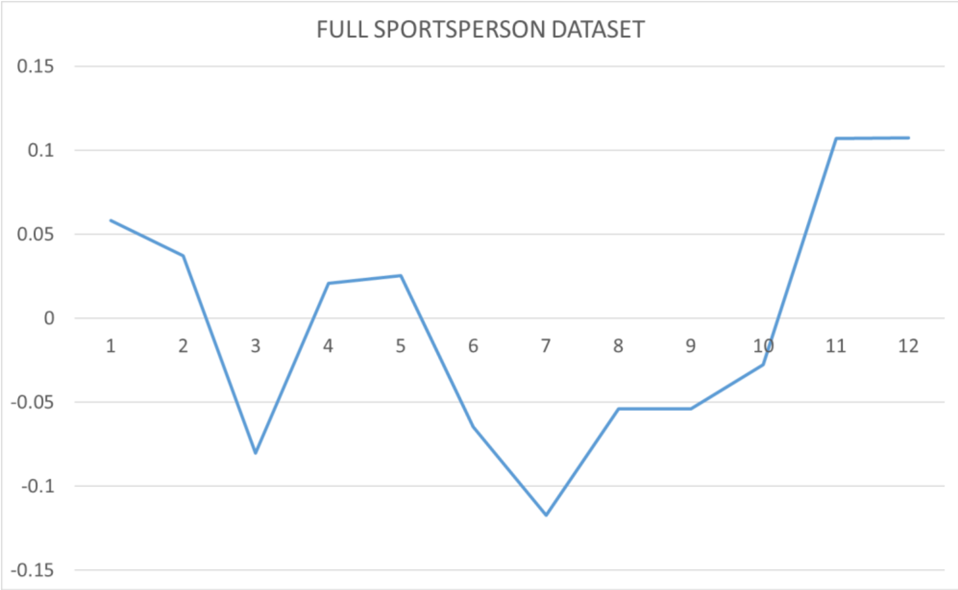
A) The baseline is calculated from the total dataset posted (for the "expected" set), rather than the national dataset M. Gauquelin used.

B) M. Gauquelin's sample comprised French sportspersons only. This effectively adds a second category to his dataset, whereas the present sample comprises many nationalities—so the sample used is assembled according to different criteria than the original. Which at least will excuse any slight divergence—we won't allow a strong divergence.

C) I have plotted only angular separation between Mars and Ascendant (so Equal Houses) rather than the time-based sectors of Gauquelin's study.

Even so, it can be seen that the original results are *broadly* replicated—at a 1% significance level (stronger than the 5% quoted in Wikipedia) there is indeed a difference between the observed and expected distribution, and this difference in particular comprises a bulge over the arc around Mars' just having risen. In fact, all the arc comprising the Eleventh to the Second House seems significant. We also have— well, I'd be generous to call it a peak, it's negative!— but an *upward tending* bump at the 9th House. But it's on the shape of the pattern that the BBC challenges the data, and by this criterion, the pattern is indeed reproduced. So I don't have to invoke the let-out clause that this dataset was assembled under different criteria to Gauquelin's.

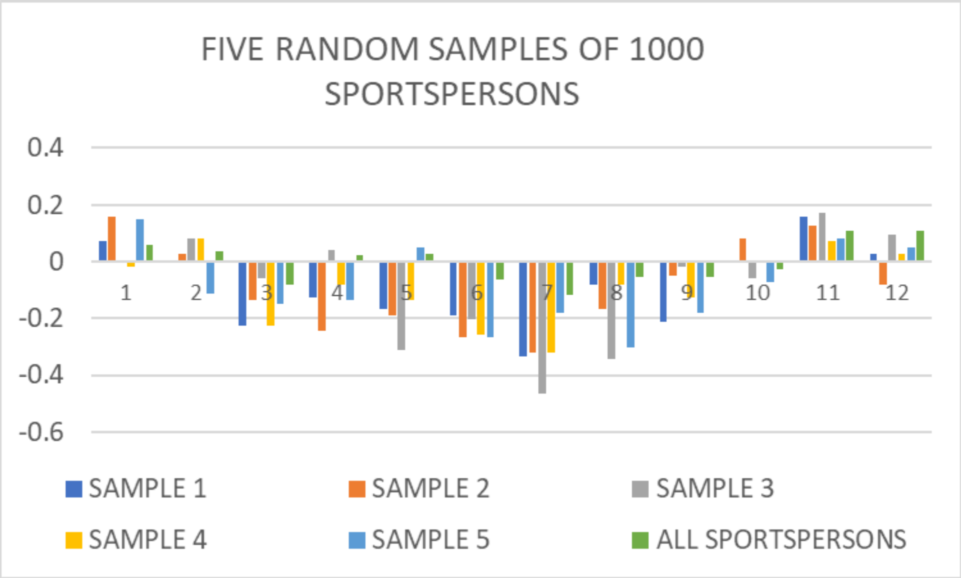




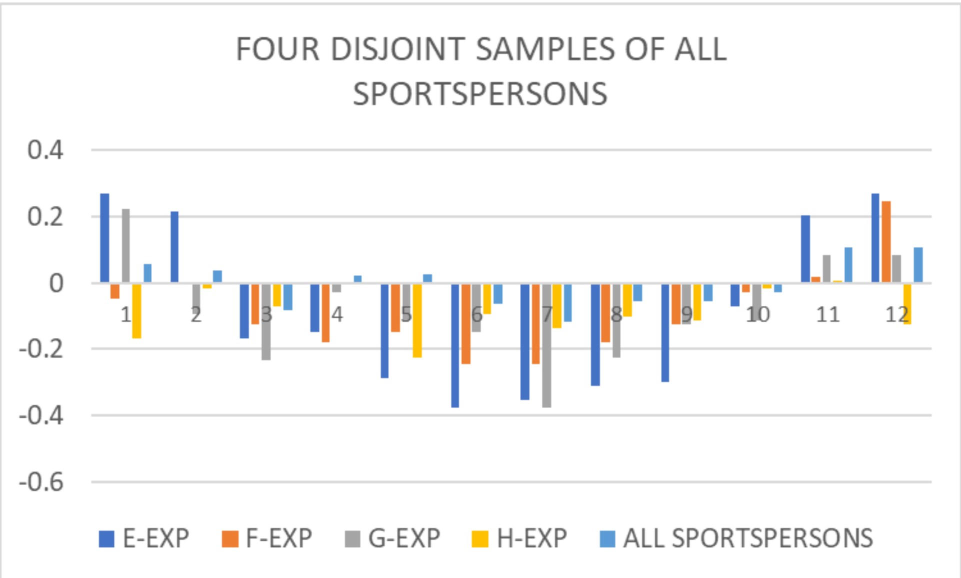
Classes: 1, 0-30°: 2, 30-60°: 3, 60-90°: 4, 90-120°: 5, 120-150°: 6, 150-180°: 7, 180-210°: 8, 210-240°: 9, 240-270°: 10, 270-300°: 11, 300-330°: 12, 330-360°:

Still, Point 5 now becomes more important: using only this dataset, can we show that the pattern runs through every part of the data?

Below, you can see a stacked line chart of the difference between observed and expected distributions of Mars in five random samples of 1000 from the Sportsman set, alongside the equivalent from the complete experiment above. The distributions certainly follow the same pattern—in most categories, either all samples show negative frequency or all show positive frequency, with the only strong divergence being in the low-scoring categories. More important than this, however, the patterns, or trends (bulge and dip!) are clearly similar. This refutes the BBC's point. As to the argument of multiple comparisons, the involvement of Mars is still what an Astrologer would expect—so this argues against chance.



And here are four disjoint sets of 1000 from the same dataset. This is stronger confirmation, in that we effectively have four independent datasets. Again, the frequencies agree in most categories, and we have that same bulge and dip.



The same pattern runs through all the data!

The BBC pundit's statement, then, is not correct as it stands. If the "multiple comparisons" argument is intended, then it's a stronger objection, but we can argue against it on sound statistical grounds. (Though actually the statement isn't even accurate, since Gauquelin used not only 12, but 18 categories or more—I'm the one who finds 12 convenient).