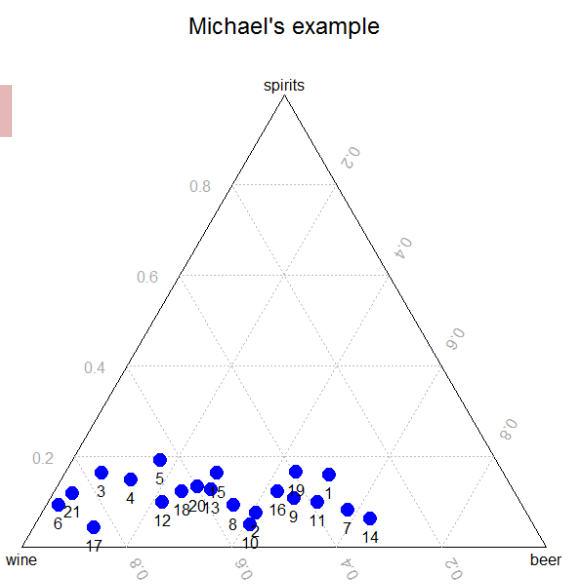
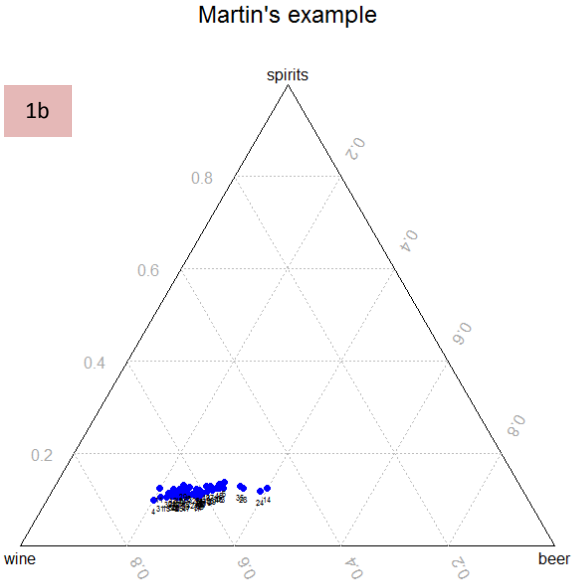
Thanks to Martin for taking the time to supply this interesting counter-example! I won't reply to his detailed rebuttals made to my previous contribution, which is becoming exhausting, probably for both of us. I can only agree to disagree with almost everything he says! But the new counter-example is interesting, so here is a side-by-side comparison of our respective counter-examples (see interpretations on next page):

Left hand side: visualizations/analyses of my data **Right hand side: visualizations/analyses of Martin's data**

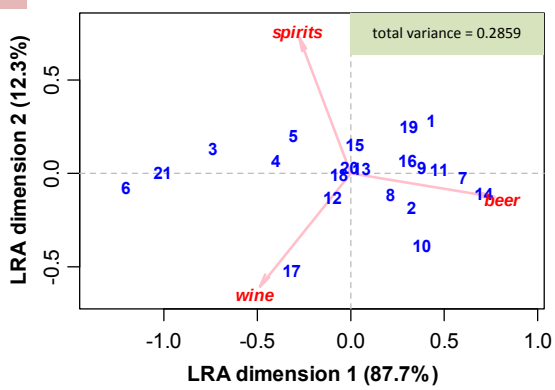
1a



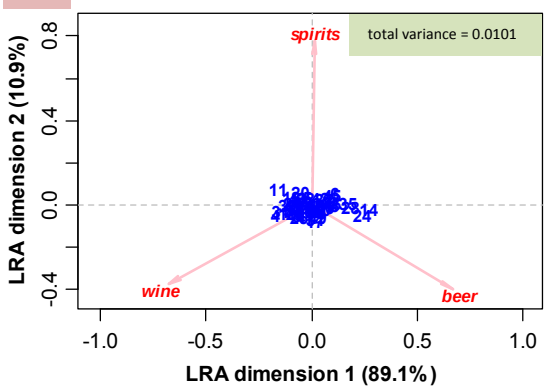
1b



2a

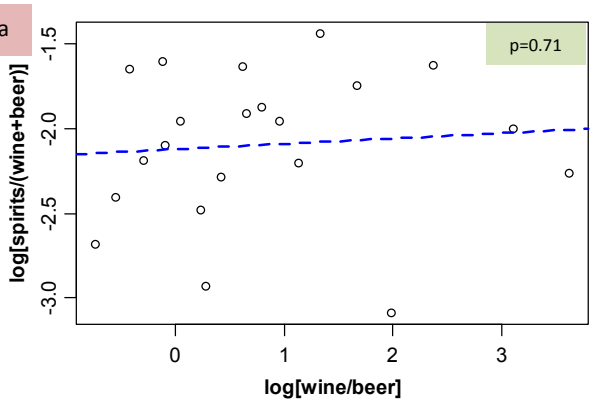


2b

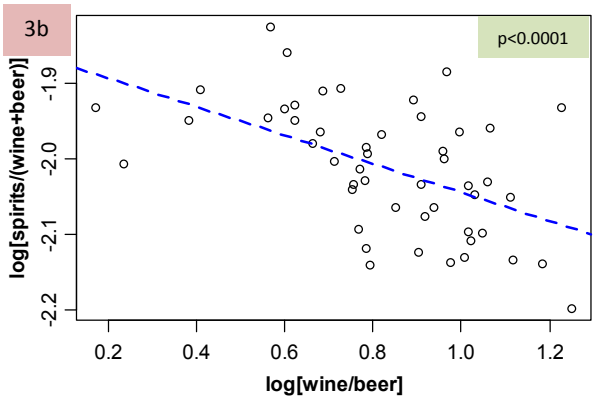


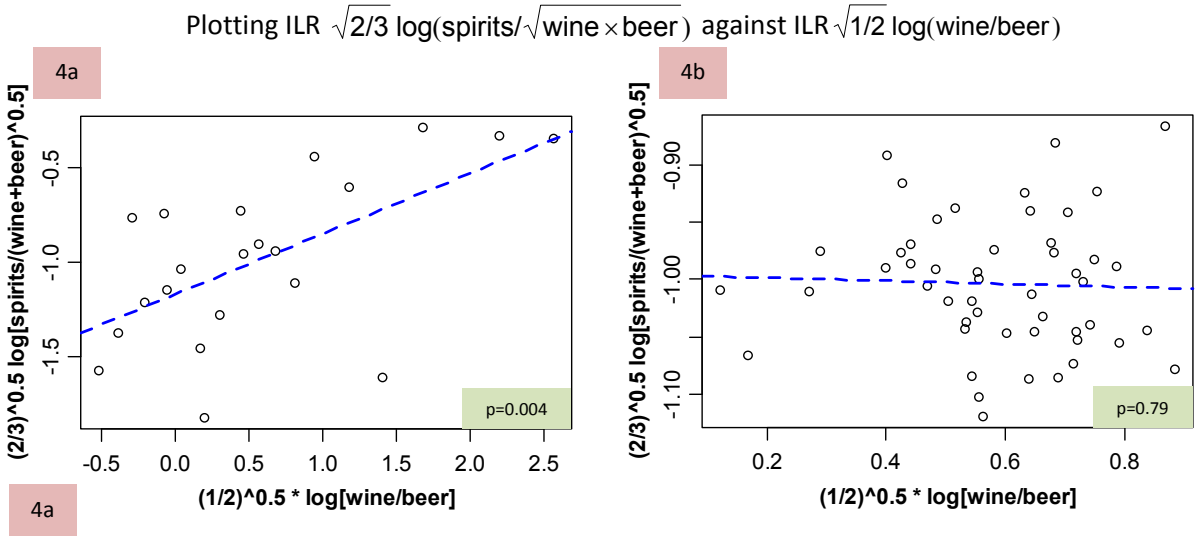
Plotting amalgamation logratio of spirits / (wine+beer) against logratio of wine/beer

3a



3b





SUMMARY: A TALE OF TWO COUNTER-EXAMPLES OF TWO LOGRATIOS

Figs 1a, b: My data has much higher variance than Martin's data, but they both show similar interchange between wine and beer, although Martin's "interchange" is over a smaller range. There is a possible side issue about signal and noise in this data set (and in mine too, perhaps my second dimension is compatible with random variation...), but that's another story. But notice the descent of his points in **Fig.1b** as one goes from right to left (increasing ratio of wine/beer)... see below!

Figs 2a, b: I did the unweighted logratio analysis, which is what the Girona group maintains, in spite of varying relative error on parts[1] The two results in **Figs 2a** and **2b** are similar with respect to the three parts, just the differences in variabilities is apparent.

We each defend our counter-examples from our point of view.

I say that in my data there is no relationship between the logratio of spirits / (wine+beer) (this amalgamation is almost constant) and the logratio of wine / beer (**Fig. 3a**), whereas the ILR creates a relationship (**Fig. 4a**), entirely due to the geometric mean of wine and beer varying in the denominator of the ILR, which is of course related to log(wine/beer).

Martin says that in his example there is no relationship between the ILR, i.e. logratio of spirits / geometric mean of wine and beer and logratio of wine/beer (**Fig. 4b**), but the amalgamation logratio creates a relationship (**Fig. 3b**).

My reply is the following. The researcher is (please remember!) interested in the contrast between high alcohol drinks and low alcohol drinks. She amalgamates the wine and beer and computes the logratio between spirits and wine+beer, which is the same as the "log-odds" log(spirits/not spirits), a quantity that she understands perfectly. To take into account the contrast between the two low alcoholic drinks summed in the amalgamation, she computes the logratio of wine/beer. All of this is pretty straightforward and logical!

In Michael's example she concludes that as the wine/beer ratio increases, there is no increase in the spirits/not spirits ratio. In Martin's example, she concludes that for a relative increase in the wine/beer ratio, there is a relative decrease in the spirits/not spirits ratio, which is quite easy to quantify, but is small. **This can, in fact, be seen in Fig.1b:** As wine/beer ratio increases (going from right to left in **Fig.1b**), the spirits to non-spirits ratio **is** decreasing. That is, people who tend to prefer wine over beer tend to drink less high alcohol spirits compared to the lower alcohol wine&beer (not spirits), which is exactly what is shown in **Fig.3b**! This effect does not exist in Michael's data (the scatterplot in **Fig.3a** is pretty flat). But Martin's ILR plot in **Fig.4b** cancels out this interesting effect in his own data.

So when our researcher also tries the ILR approach, she gets diametrically opposite results. She tries her best to understand what an ILR is, what the strange constants are and why the geometric mean in the denominator of the ratio changes the results so much compared to the simple amalgamation. To try to enlighten herself, she plots the amalgamations against the geometric means in each example, sees quite a close agreement, so is even more puzzled.

---

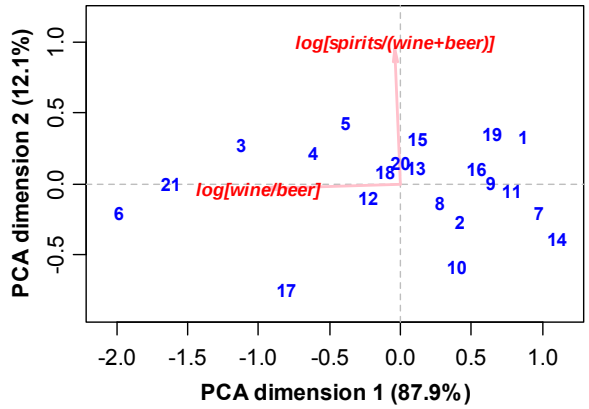
[1] Authors like Filzmoser and Hron have started to consider weighted analyses, which is what I have been advocating for about 15 years since my collaboration with Paul Lewi. In the 1980s, simultaneously with the appearance of Aitchison's unweighted multivariate approach, Paul independently invented the spectral map, which is what I call weighted logratio analysis. Lewi was ahead of his time (see https://en.wikipedia.org/wiki/Paul_Lewi), yet – to my knowledge – his contribution has never been acknowledged by the CoDa group. Also see: Lewi P.J. (1989). "Spectral Map Analysis. Factorial analysis of contrasts, especially from log ratios." In: *Chemometrics and Intelligent Laboratory Systems,* 5 (2), 105-116.

My conclusion and advice to the researcher: In my data the ILR creates a relationship that doesn't exist (and I created the data with this non-relationship!), and in Martin's data the ILR hides a relationship that can be seen in the data and is statistically significant. I advise the researcher to totally forget about the ILR, and use the amalgamation logratio, which fits perfectly the objective of her research. With relief, she settles with the results that use the straightforward logratio amalgamation, which she understands and can interpret.

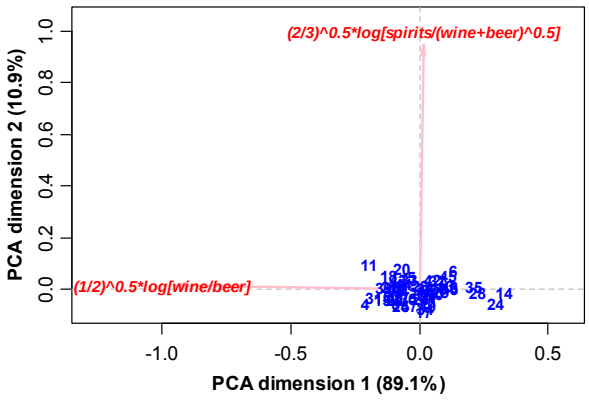
P.S. Here below are the (unweighted & unstandardized) PCAs of the two-column data matrices involving the pairs of logratios analysed in the two data sets. The ILR analysis in **Fig.5b** is identical to the logratio analysis in **Fig.2b**, which is the much vaunted theoretical benefit of this approach. The simpler approach in **Fig.5a** almost exactly (Procrustes correlation = 0.9903) reproduces the configuration in the logratio analysis of **Fig.2a**, and offers the practical benefit of having variables that are much easier to interpret (and don't need funny constants!). So Martin & co. are effectively criticizing the amalgamation approach here because of the < 0.001 recovery of the "true" inter-sample distances. In many examples I have done so far, on much bigger data sets, I am finding that amalgamation logratios approximate the logratio geometry very closely, so they are preferable to ILRs if one is prepared to take this simpler, and in many cases more substantively applicable, approach, which John Aitchison himself so wisely recommended many years ago).

Notice that, since the two variables in each PCA below are almost exactly identified with the two principal axes, these two PCA biplots confirm the conclusions in **Fig.3a** and **Fig.4b** respectively.

5a with Michael's simple ratios (cf. Fig.2a)

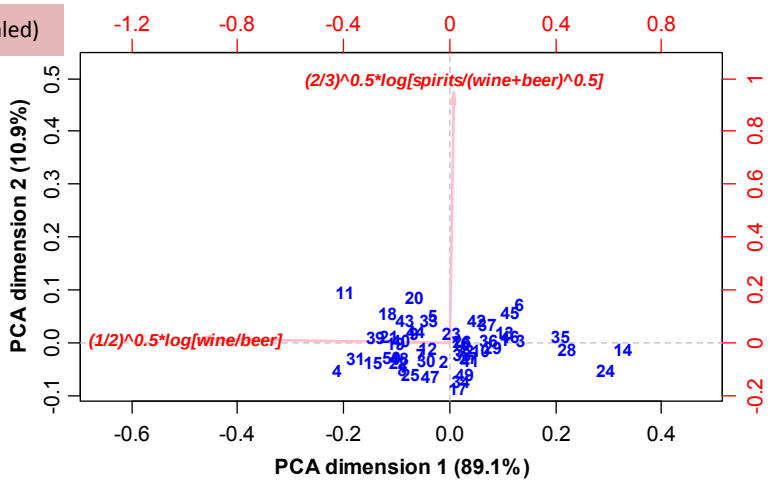


5b with Martin's ILRs (cf. Fig.2b)



To enlarge the low-variance sample points in the PCA of Martin's ILRs on the right, here is an alternative biplot where samples are still shown on the same scale, but the two ILRs are shown with a different scale (red scale):

5b (rescaled)



Michael Greenacre
Grèixer
17 March 2018