

TheTrophicLink

{ 2014.04.05 }

About p-values

I stepped into a blog post by Pia Parolin titled “Do all biological processes need to be statistically significant?” (<http://tinyurl.com/nweq5xa>). It sounds, at moments, the frustrated cry of the field biologist observing cool patterns and building cool theories on it until he/she faces that bloody $p\text{-value}=0.051$. Who hasn't been there? Yet Pia's article contains more than that, and it raises interesting issues (also see the article's comments). Here are some notes of mine.

Modern-day ecology is still flawed by many “Old-School Observations” that somehow became paradigms, “articles of faith” because they made up a cool story without an appropriate statistical testing ever being carried out. Jeremy Fox calls them “zombie ideas” and he wrote many interesting blog posts and one cool paper about it: see [here](#) and [here](#).

Obviously it is also true that statistics without observation and knowledge of the system can be pointless. Say that we are studying the change in generation time of an organism across a temperature gradient: the analyses give a “very significant” result, but the change corresponds to some minimal fraction of the organism's life span. Our result would be significant statistically, but probably not biologically.

En passant, I would like to point out that statisticians often criticise biologists' habit of adding an adverb to “significant”: they intend the p -value as a 0/1 threshold, so the result is either significant or non significant, never “almost significant” or “very significant” ([here](#) are some guidelines that a statistician I know would find rather amusing). After all, we either refuse a null hypothesis or not, we can't “almost refuse” it.

Statistics might seem dry and for sure has limitations, but it is a tool to restrict human attitude toward interpretation. Interpreting a pattern without the tether of stats would leave dangerous room for subjectivity and thus, for example, to retain a cool theory as true even without enough data to support it.

I have just found [an article on Nature on statistical errors](#), which links to a special issue on “Challenges in irreproducible research”: it might be of interest, I'll go through it as soon as possible. For further reading, ~~this month's issue of ECOLOGY~~ hosts an entire Special Forum on P Values and Model Selection.

PS (practical note): interpreting statistical tests only in the light of p -values is a limiting view. Statistics is obviously much more and p -values are a tool that should be used together with others, such as R-squared (representing the amount of variance in the data explained by the model) and confidence intervals (perhaps more intuitive to interpret than p -values and particularly useful when the estimated p -values are not universally deemed reliable, as in Mixed Effect models).

[originally published on www.marcoplebani.com]

Like 0



[Edit this entry.](#)

Posted by [Marco Plebani](#) on [Saturday, April 5, 2014, at 8:49 pm.](#)

Filed under [Uncategorized](#). Follow any responses to this post with its [comments RSS](#) feed. Comments are closed, but you can [trackback](#) from your blog.