Significance tests in climate science

Ambaum, Maarten HP. "Significance tests in climate science"
Journal of Climate (2010)

Problem

- “this correlation is 95% significant”
- “the significant values are printed in bold”
- “the trend is significant and therefore likely physical”

Problem

- In a recent issue of a top journal in the field, 3/4 of articles showed misuse of significance tests
- Probably even worse in some other fields?

Spot the error

(i) My measurement stands out from the noise.

(ii) So, my measurement is not likely to be caused by noise.

(iii) It is therefore unlikely that what I am seeing is noise.

(iv) The measurement is therefore positive evidence that there is really something happening.

(v) This provides evidence for my theory.

General confused scientist thought process

(i) My measurement stands out from the noise.

(ii) So, my measurement is not likely to be caused by noise. **$p(measurement|H_0)$ is low**

(iii) It is therefore unlikely that what I am seeing is noise. **$p(H_0)$ is low**

(iv) The measurement is therefore positive evidence that there is really something happening.

(v) This provides evidence for my theory

this makes $p(H_1)$ more likely

Example: correlated time series

- We measure some strong correlation r_0 between two time series
- H: the time series are related, r_0 is a measure of this relation
- H_0: they are not related, r_0 is a fluke
- Q: how likely is H0 to be valid? Lets do a significance test!

Example: correlated time series

- Q: how likely is H_0 to be valid? Lets do a significance test!
- ...well, no, a significance test *done properly* will give us $p(r>r_0|H_0)$, not $p(H_0|r>r_0)$
- = **error of the transposed conditional**

TABLE 1. Example frequency table for a typical test of significance of a correlation.

	Low r	High r
Related	40	60
Unrelated	95	5

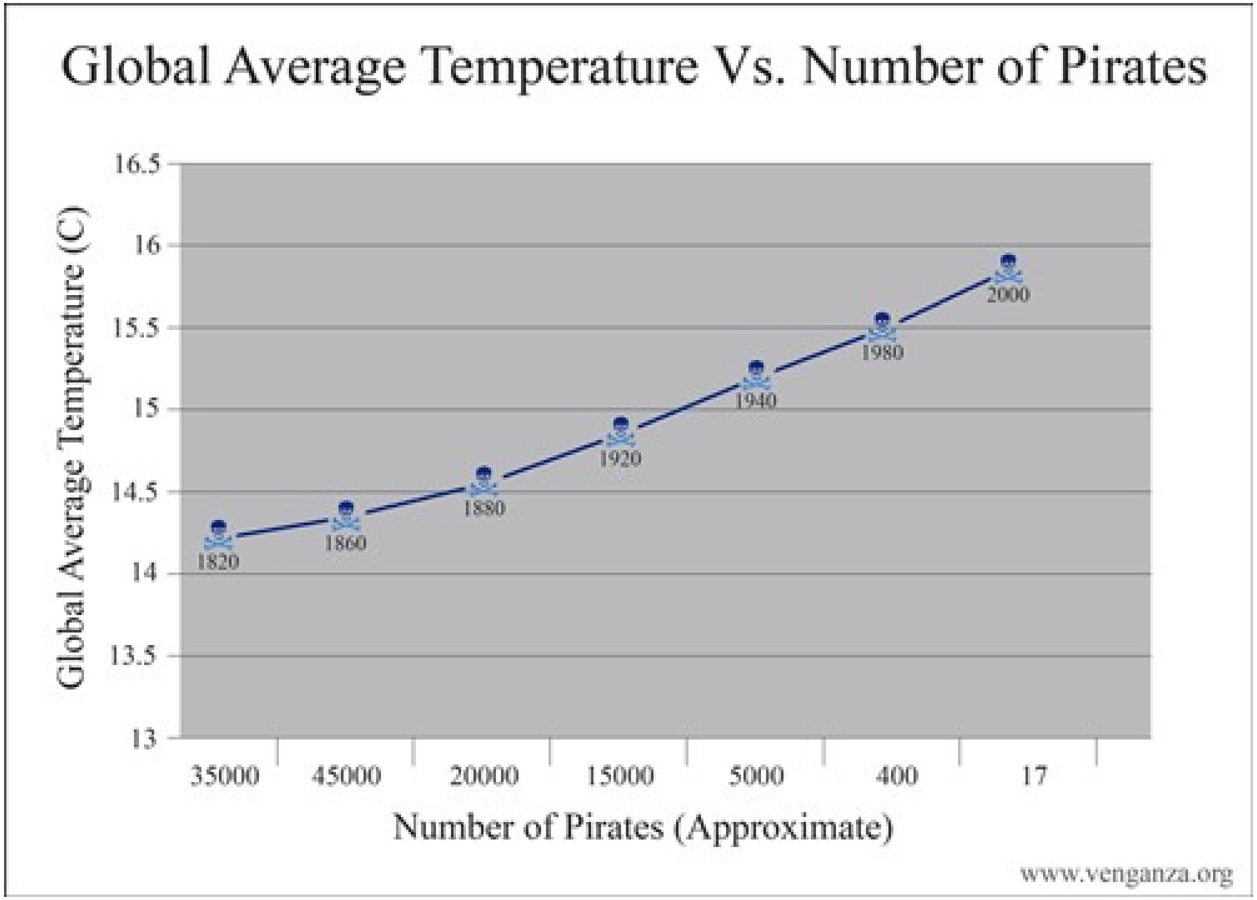
TABLE 1. Example frequency table for a typical test of significance of a correlation.

	Low r	High r
Related	40	60
Unrelated	95	5

- *“the time series are significantly correlated (p=0.05) ”* = category error, since the low p value is a property of the unrelated time series, not related time series

- So what is $p(H|r>r_0)$,?
- Bayes theorem:

$$p(H|r > r_p) = 1 - p(r > r_p|\bar{H})\frac{p(\bar{H})}{p(r>r_0)},$$


Global Average Temperature Vs. Number of Pirates
Global Average Temperature (C)
16.5
16
15.5
15
14.5
14
13.5
13
1820
1860
1880
1920
1940
1980
2000
35000
45000
20000
15000
5000
400
17
Number of Pirates (Approximate)
www.venganza.org