Null and alternative hypotheses

Jeff Leek

@jtleek

www.jtleek.com

Hypothesis testing

- A form of inference
- A decision rule

H_0: The relationship between age and expression is exactly zero

H_0: The relationship between age and expression is not zero

$$Expr = b_0 + b_1 \, Age + e$$

H_0: $b_1 = 0$

H_0: $b_1 \neq 0$

Characteristics of hypotheses

- Not possible to accept the null
- Ideally your statistic is set up to be "monotone" so bigger/smaller is "less null"
- Null changes if adjustments change
- Null must make intuitive sense
- This is very important to get right

Hypothesis testing is controversial

- The null matters
- The null may always be false

Pachter's P-value Prize

May 26, 2015 in reviews | Tags: Bruce Birren, Eric Lander, evolution, Manolis Kellis, neofunctionalization, p-value, prize, S. cerevisiae, subfunctionalization, yeast | by Lior Pachter | 140 comments

In this blog post **I offer a cash prize** for computing a *p*-value [update June 9th: the winner has been announced!]. For details about the competition you can skip directly to the challenge. But context is important:

Background

I've recently been reading a bioRxiv posting by X. Lan and J. Pritchard, Long-term survival of duplicate genes despite absence of subfunctionalized expression (2015) that examines the question of whether gene expression data (from human and mouse tissues) supports a model of duplicate preservation by subfunctionalization.

The term *subfunctionalization* is a hypothesis for explaining the ubiquity of persistence of gene duplicates in extant genomes. The idea is that gene pairs arising from a duplication event evolve, via neutral mutation, different functions that are distinct from their common ancestral gene, yet together recapitulate the original function. It was introduced in 1999 an alternative to the older hypothesis of *neofunctionalization*, which posits that novel gene functions arise by virtue of "retention" of one copy of a gene after duplication, while the other copy morphs

Notes and further reading

- Inference class:
  - https://www.coursera.org/course/statinference
- Statistics and R for the Life Sciences
  - https://www.edx.org/course/statistics-r-life-sciences-harvardx-ph525-1x