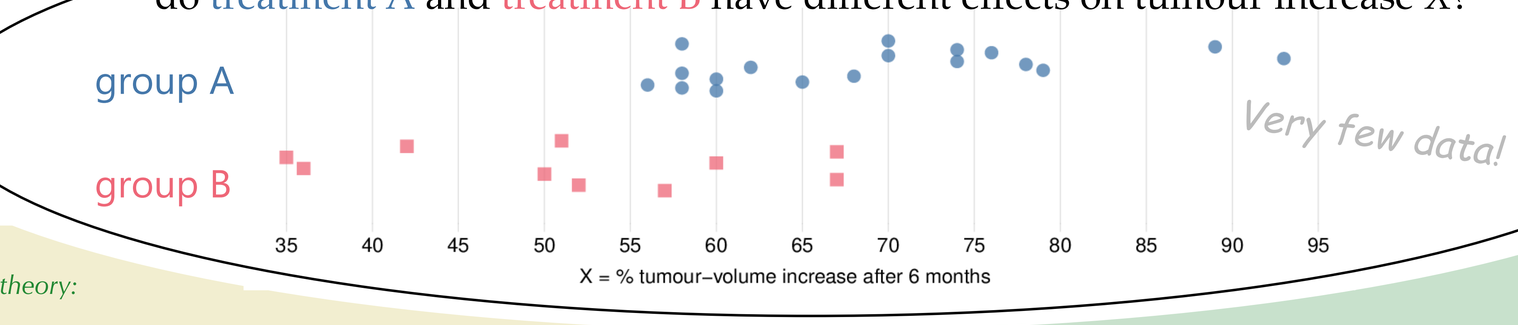
USER-FRIENDLY SOFTWARE FOR BAYESIAN ANALYSIS OF MEDICAL DATA

PierGianLuca Porta Mana, Kjetil Dyrland, Alexander Selvikvåg Lundervold

Dept of Computer Science, Electrical Engineering, and Mathematical Sciences, HVL; MMIV, HUS

Sampling theory

Simple example question:
do treatment A and treatment B have different effects on tumour increase X?



Bayesian theory

You are probably familiar with the typical conclusions from sampling theory:

" Assuming that the two full populations:
- are gaussian and independent
- have identical variances (F-test, $p = 0.44$)

*It's actually slightly more probable that the variances are **different***

and that the sample sizes where decided before running the experiment, then:

The hypothesis that the population means are equal has a p-value 0.00043 (two-tail t-test: +4.0)

What does the number "0.00043" really mean?

*This p-value is wrong If the sample sizes 18 & 10 were **decided not in advance** but by some other rule!*

The sample mean difference of X is17.6
The 95% confidence interval is [8.6, 26.7] "

*This does **not** mean that $8.6 < X < 26.7$ with 95% probability! It means that this technique to construct the interval contains the true value in 95% of all imaginary datasets. But what's the best interval that contains X for **our** dataset?*

American Statistical Association's official statement

THE AMERICAN STATISTICIAN
2019, VOL. 73, NO. S1, 1–19: Editorial
https://doi.org/10.1080/00031305.2019.1583913

Moving to a World Beyond "$p < 0.05$"

- Don't base your conclusions solely on whether an association or effect was found to be "statistically significant" (i.e., the p-value passed some arbitrary threshold such as $p < 0.05$).
- Don't believe that an association or effect exists just because it was statistically significant.
- Don't believe that an association or effect is absent just because it was not statistically significant.
- Don't believe that your p-value gives the probability that chance alone produced the observed association or effect or the probability that your test hypothesis is true.
- Don't conclude anything about scientific or practical importance based on statistical significance (or lack thereof).

2. Don't Say "Statistically Significant"

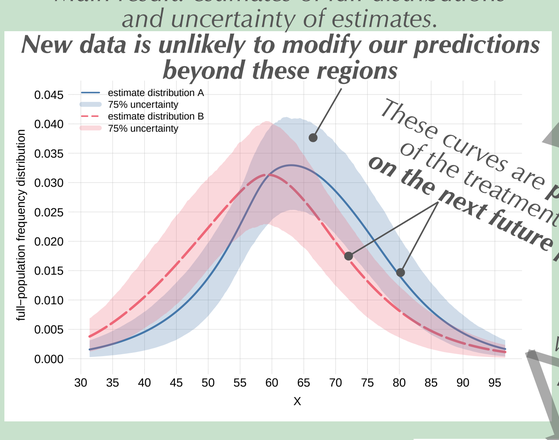
We conclude, based on our review of the articles in this special issue and the broader literature, that it is time to stop using the term "statistically significant" entirely. Nor should variants such as "significantly different," "$p < 0.05$," and "nonsignificant" survive, whether expressed in words, by asterisks in a table, or in some other way.

4. Other Approaches

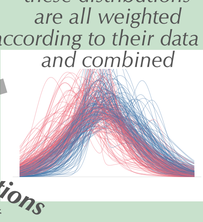
In view of the prevalent misuses of and misconceptions concerning p-values, some statisticians prefer to supplement or even replace p-values with other approaches. These include methods that emphasize estimation over testing, such as confidence, credibility, or prediction intervals; Bayesian methods; alterna-

An intuitive understanding of how Bayesian probability theory works:

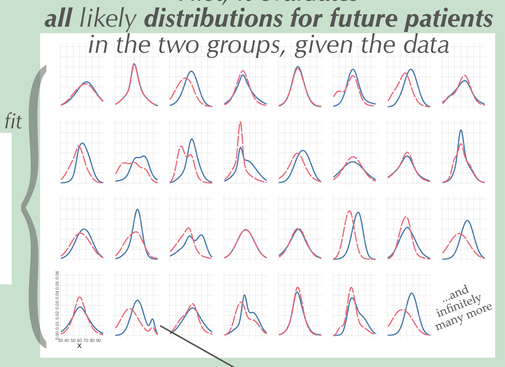
Main result: estimates of full distributions and uncertainty of estimates.



these distributions are all weighted according to their data fit and combined

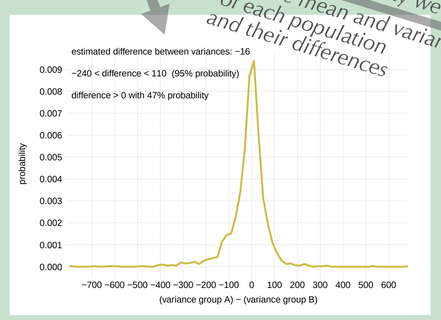


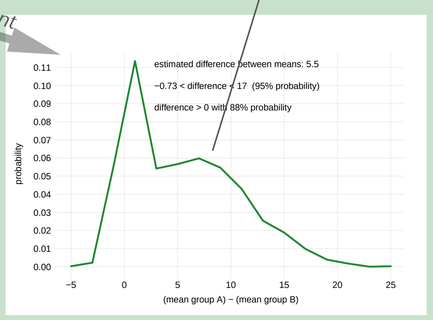
*First, it evaluates **all** likely **distributions for future patients** in the two groups, given the data*



No assumptions of gaussianity anywhere

We can estimate any quantity we want For example the mean and variance of each population and their differences





*These detailed probabilities can be used in a **cost/benefit analysis** (decision theory)*

"Treatment A costs \$, gives a% benefit. Treatment B costs \$\$, gives b% benefit. Which treatment should we choose?"

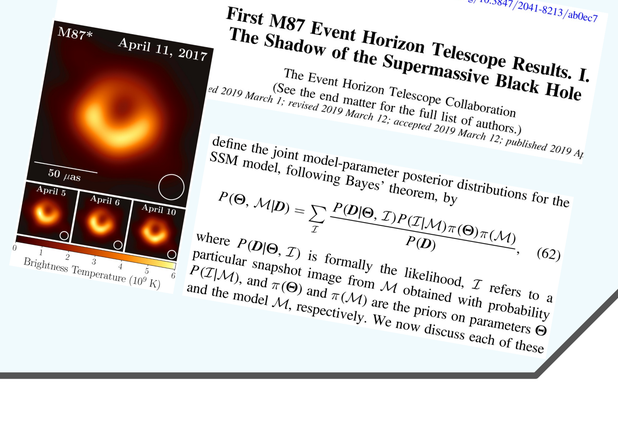
" Assuming that the two full-population distributions are probably smooth, we predict:
- The distributions of **future treatment outcomes** will be as in the plots, within uncertainties shown
- Future patients under treatment A will have on average $X = 65$, and $60 < X < 71$ with 95% probability
- Future patients under treatment B will have on average $X = 59$, and $51 < X < 65$ with 95% probability
- The difference between the treatment means will be within [–0.73, 17] with 95% probability
- Average X under treatment A will be **larger** than under treatment B with 88% probability
- Variance of X under treatment A will be **smaller** than under treatment B with 53% probability "

Compare the two analyses

The statements of sampling theory are:
- vague
- obscure or misleading ("significant"?)
- heavily dependent on tacit assumptions
  gaussianity, stopping rules, ...
- subject to doubts and corrections
  "should I use this test or that test?" "Bonferroni correction?"

The predictions of Bayesian theory are:
- detailed
- quantitative
- easy to interpret and understand
  eg: "fraction x% of population will have effect y"
- calculated always in the same way
  no matter the sample size, no matter how many hypotheses

Bayesian theory is used, for example, for all inferences about the black hole observed in the galaxy M87



This was just a simple example. Bayesian theory deals in the same way with multiple hypotheses, variates, and correlation questions

True! That's why we are developing a user-friendly app to do Bayesian analysis on (non-imaging) medical data

Machine Learning

The maths will be taken care of under the hood
The software will suggest meaningful questions to be asked (in line with ASA's statement)
Works with mixtures of continuous, integer, categorical variables
– can make predictions about their correlations and relevance
Automatic imputation
No assumptions of gaussianity, linearity or other (ie: nonparametric)
No need to wonder which test or formula to use, no corrections of any kind

We're already using a prototype version for drug-discovery and Alzheimer research

Please get in touch if you want to test it and help us making a great sofware!

Bibliography

American Statistical Association 2016: *Statement on Statistical Significance and p-values*
American Statistical Association 2019: *Moving to a World Beyond "p < 0.05"*
Jaynes 2003: *Probability Theory: The Logic of Science*
MacKay 1994: *Information Theory, Inference, and Learning Algorithms*
Hjort, Holmes, Müller, Walker 2003: *Bayesian Nonparametrics*
Müller, Rodríguez 2013: *Nonparametric Bayesian Inference*
Dunson, Bhattacharya 2011: *Nonparametric Bayes regression and classification through mixtures of product kernels*
Sox, Higgins, Owens 2013: *Medical Decision Making*
The Event Horizon Telescope Collaboration 2019: *First M87 Event Horizon Telescope results. VI. The shadow and mass of the central black hole*