



Gorgias and the Selection of Statistical Methodologies

Nikos Asimakis,
Yiannis Kiouvrekis,
Petros Stefanias
December 19, 2018

National Technical University of Athens, Department of Mathematics

Table of contents

1. Introduction
2. Problem Recognition and Specification
3. System Design
4. System Implementation
5. Chi square test implementation
6. Review and Maintenance
7. Conclusions

Introduction

- Problem: Errors in statistical analysis
 - Wide use of statistics in different fields (e.g. medical research)
 - Researchers are often not specialists in mathematical statistics
 - Wrong methods/wrong interpretation of the results
- Solution: An information system based on non-monotonic logic
 - use Gorgias argumentation based framework
 - augment the use of already existing software packages (e.g. R)
 - develop an agent to combat common errors in statistical analysis
 - choose the correct method based on the problem and the sample
 - interpret the result, like a mathematician, and explain the answer with references to rules and conditions

Problem Recognition and Specification

Problem Recognition and Specification

- The core of the particular ecosystem is human behavior and the way we can approach it
- Mathematicians have developed numerous theories about modelling human systems
- Why do we need so many mathematical theories? Is there no universal theory to describe human groups?
- The problem of the *irrational* behaviour
- The corresponding phenomenon occurs when the researcher is asked to make a decision about the model he/she needs to use in the data analysis. Although all modern statistical and data analysis software contains all the information the researcher needs, errors are common.

Types of errors

- Errors appear at the initial stages, such as during the literature review in order to specify the basic research questions.
- During the important statistical/mathematical stages we see errors related to p-values (p-values given in closed form, p-values missing after statistical tests, incorrect p-values, incorrect demonstrations of p-values), errors related to statistical tests (insufficient data presented for the statistical test, incorrect name of the statistical test, use of incorrect test, lack of normality), failure to summarize the data, failure to demonstrate the findings mathematically, and errors in using statistical symbols

System Design

- In order to be able to manage the contradictions of human behaviour, we need a logical system that can handle the concept of contradiction. Here we have to be very careful because in classical logical systems, such as propositional logic and first order logic, if a proposition theory contains a contradiction, it is then inconsistent and thus uninteresting.
- Based on the above, it is natural to use a system like Logic Programming without Negation as Failure (LPwNF)

In LPwNF logic programs are nonmonotonic theories where each logic program is a pool of default sentences from which we select a suitable subset, called extension, to reason with.

Sentences are written in the usual logic programming language. We will refer to these sentences as rules.

Rules are assigned locally a relative strength through a partial ordering relation

Example 3.1

Consider the follow program - set of rules :

$$\text{bird}(x) \rightarrow \text{fly}(x) \quad (1)$$

$$\text{penguin}(x) \rightarrow \neg\text{fly}(x) \quad (2)$$

$$\text{penguin}(x) \rightarrow \text{bird}(x) \quad (3)$$

$$\text{bird}(\text{Tweety}) \quad (4)$$

with an ordering relation $r_2 > r_1$

From this set of rules we can conclude that $\text{fly}(\text{Tweety})$. If we add the $\text{penguin}(\text{Tweety})$ as a statement then we can take $\text{fly}(\text{Tweety})$ and $\neg\text{fly}(\text{Tweety})$ from the rules of the program, but due to the ordering relation the second conclusion overrides the first.

In essence, we described mathematically what everyone understands: although a bird usually flies, the penguin, which is a bird, does not fly.

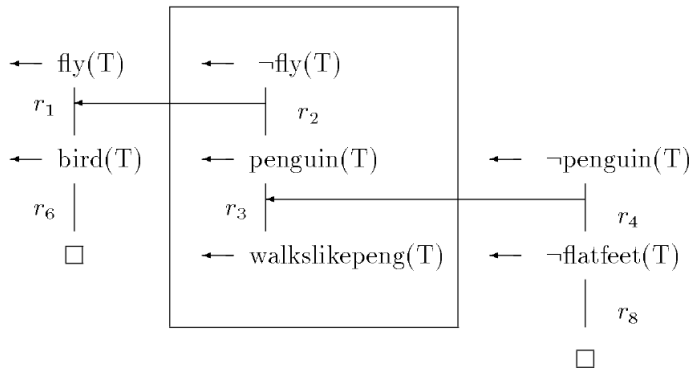
Proof procedure in LPwNF is based on the attack relation between rules and the construction of an admissible subset of rules, e.g.:

$$\begin{aligned}r_1 &: \textit{fly}(x) \leftarrow \textit{bird}(x) \\r_2 &: \neg \textit{fly}(x) \leftarrow \textit{penguin}(x) \\r_3 &: \textit{penguin}(x) \leftarrow \textit{walkslikepeng}(x) \\r_4 &: \neg \textit{penguin}(x) \leftarrow \neg \textit{flatfeet}(x) \\r_5 &: \textit{bird}(x) \leftarrow \textit{penguin}(x) \\r_6 &: \textit{bird}(t) \\r_7 &: \textit{walkslikepeng}(t) \\r_8 &: \neg \textit{flatfeet}(t)\end{aligned}$$

where $r_2 > r_1$ and $r_4 > r_3$

Logic

Proof that $fly(t)$:



System Implementation

- Gorgias is a general framework of argumentation theory that combines its ideas of preference reasoning and abduction in a way that retains both assets. It can form the basis for reasoning with adaptable preference policies in dynamic and evolving environments, despite incomplete information;
- Gorgias' syntax is based on Prolog;

Representation of Knowledge and Belief

To say something flies when it is a bird and that something does not fly when it is penguin we write:

```
rule(r1(X), fly(X), [bird(X)]).  
rule(r2(X), neg(fly(X)), [penguin(X)]).
```

In this example, it is clear that these two rules are in conflict when something is a penguin and a bird:

```
rule(f1, bird(tweety), []).  
rule(f2, penguin(tweety), []).
```

To solve the conflict, we use the special hypothesis **prefer** / 2. Therefore, for in our example we have:

```
rule(pr1(X), prefer(r2(x), r1(x)), []).
```


Chi square test implementation

We will use the above to describe our basic methodology for developing an information system for data analysis and statistical analysis using that uses argumentation logic.

For calculating the statistical values we will use the language R. Our information system will choose the correct statistical control, depending on the characteristics of the sample and if the theorem conditions are satisfied, and then it will use Gorgias to interpret these results and perform hypothesis testing.

This architecture is chosen because:

- R is the basic scientific data analysis tool and provides many ready-made statistical functions, so we do not have to implement them, and we can also take advantage of the fact that they are tested because of the widespread use of R in the field of the statistical analysis.
- The logic for the interpretation of the results is concentrated in a Gorgias rule file, and it fully corresponds to what is reported in statistics literature. This way we are sure of the correctness of our interpretation. In addition, there are development benefits because the user who knows statistics can only deal with drafting these rules and not with the rest of the application.

Also, if in the future these rules need to be corrected or extended, then the code for the rest of the program does not need to be changed, as long as the interface between different parts of the program remain stable.

R and Gorgias

- The use of Gorgias gives us more possibilities than Prolog. For example we can more accurately model hypothesis testing as described in the literature by making use of the ordering relation between rules as described in previous sections.
- Furthermore we can take advantage of Gorgia's ability to present us with the admissible argument constructed during the proof procedure and thus "justify" its answers. This way, along with the result, the user will be informed on which statistical method was used and why it was chosen. As we will see in more detail below, if the conditions of the central boundary theorem are satisfied, X^2 will be used for control, otherwise another method will be used and the user will be notified accordingly.
- The previous procedure enables the user to check the correctness of the result, since the logic of the program is not be hidden inside it but will be/can be shown when the program is running. Finally, the user, who is often not a mathematician, will "learn" (be trained in) the correct use of statistics.

Review and Maintenance

Review and maintenance.

In order to show that our system works, we will conduct a statistical control to check if there is dependence between two attributes: A and B . One way we can do this is to use a χ^2 control independent test.

To apply this test, we typically represent the data from one sample size n , in the form of a 2×2 **contingency table**. This matrix is a frequency table where we have them two subpopulations (attribute categories B) and their columns successes and failures (A attributes).

	Success	Failure	Total
First subpopulation	n_{11}	n_{12}	n_1
Second subpopulation	n_{21}	n_{22}	n_2
Total	n_1	n_2	n

Table 1: 2×2 Contingency Table

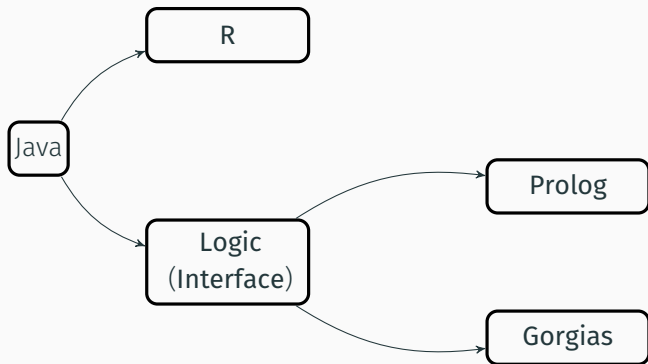


Figure 1: The basic structure of our system

```
void load(String rfile);  
void closeR();  
void chisq(String variable);  
double getChisqStat(String variable);  
int getChisqDf(String variable);  
double getChisqPValue(String variable);  
double getMinExpected(String variable);  
void fisher(String variable);  
double getFisherPValue(String variable);
```


Logic Interface

```
void load(String file);
void claim(String condition);
void claim(String condition, String label);

boolean test(String condition);
List<String> query(String variable, String condition);
List<List<String>> query(List<String> variables, String
List<List<String>> why();

List<String> listPredicates();
void disclaimAll();
void disclaimLast();
void disclaim(String condition);
String negate(String condition);
```

Case Control

The **Statistics** class is responsible for the handling of R and the collection of statistical results. Based on what we said in the introduction, interpretation of these results will be made using the functions of interface **Logic**. We will take advantage of the Gorgias's argumentation logic, which will lets us express our logic in a natural case-control study, as a mathematician does. Also, through Gorgia's proof procedure and the construction of the acceptable argument, we can inform the user what rules and data led to this answer.

The way we construct Gorgia's rules file should be consistent with how hypothesis testing is conducted in classical statistics. Null hypothesis and the alternative are not just two hypotheses we want to check on the basis of the data and decide which of the two is correct. That is why we cannot alternate them and that needs to be reflected in Gorgia's rule file and in the proof procedure. Null hypothesis is the hypothesis that we are questioning (for example, our query to know if two sample characteristics are independent), and the alternative case (e.g.the relevance to sample characteristics) is the case that we want to prove it is true.

Gorgias rule file for statistics

<pre>rule(h0_is_not_rejected(X), neg(rejecth0(X)), []).</pre>	<p>H_0 is initially not rejected. It will be rejected or not based on what is observed in the sample.</p>
<pre>rule(chisq_is_valid(X), chisq_valid(X), [chisq(X)]). rule(chisq_is_not_valid(X), neg(chisq_valid(X)), [chisq(X), chisq_minexpected(X,MinExpected), MinExpected < 5]). rule(prefer_chisq_is_not_valid(X), prefer(chisq_is_not_valid(X), chisq_is_valid(X)), []).</pre>	<p>X^2 test is valid when the central limit theorem can be applied. Practically this means that all observed frequencies are ≥ 5</p>

Gorgias rule file for statistics

<pre>rule(chisq_rejects_h0(X), rejecth0(X), [chisq(X), chisq_valid(X), chisq_pvalue(X, Pvalue), significance(Significance), Pvalue < Significance]). rule(prefer_chisq_rejects_h0(X), prefer(chisq_rejects_h0(X), h0_is_not_rejected(X)), []).</pre>	<p>χ^2 is used with $\gamma = 95\%$ confidence level or $\alpha = 5\%$ significance level</p> <p>If P-Value < α then H_0 is rejected</p>
<pre>rule(fisher_rejects_h0(X), rejecth0(X), [fisher(X), fisher_pvalue(X, Pvalue), significance(Significance), Pvalue < Significance]). rule(prefer_fisher_rejects_h0(X), prefer(fisher_rejects_h0(X), h0_is_not_rejected(X)), []).</pre>	<p>When the central limit theorem can not be applied we use Fisher's exact test</p> <p>Fisher's exact test is used $\alpha = 5\%$ significance level</p> <p>If P-Value < α then H_0 is rejected</p>

Example Usage

Cannot reject null hypotheses, data1 are independent
Why? `[[h0_is_not_rejected(data1)]]`

Null hypotheses rejected, data2 are codependent.
Why? `[[f1, f6, f5, chisq_is_valid(data2), f5, chisq_rejects_h0`

Cannot reject null hypotheses, data3 are independent
Why? `[[h0_is_not_rejected(data3)]]`

Null hypotheses rejected, data4 are codependent.
Why? `[[f1, f17, f16, fisher_rejects_h0(data4)]]`

Example Usage

Gorgias state

```
consult(stats.pl).
rule(f1,significance(0.05000000000000000044),[]).
rule(f2,chisq(data1),[]).
rule(f3,chisq_pvalue(data1,0.6570195719690067),[]).
rule(f4,chisq_minexpected(data1,13.714285714285714),[]).
rule(f5,chisq(data2),[]).
rule(f6,chisq_pvalue(data2,1.386470999319574E-10),[]).
rule(f7,chisq_minexpected(data2,43.07692307692308),[]).
rule(f9,chisq_pvalue(data3,0.22674842690343286),[]).
rule(f10,chisq_minexpected(data3,4.2),[]).
rule(f11,fisher(data3),[]).
rule(f12,fisher_pvalue(data3,0.24022012054762612),[]).
rule(f14,chisq_pvalue(data4,0.05485393990013243),[]).
rule(f15,chisq_minexpected(data4,4.125),[]).
rule(f16,fisher(data4),[]).
rule(f17,fisher_pvalue(data4,0.030221989999279542),[]).
```

Conclusions




Conclusions




- developed an information system that limits errors in the use of statistics in various fields and enhances the use of rich statistical packages like R
- by using the argumentation framework provided by LPwNF and Gorgias we express the logic rules that drive the program in a concise and precise way that correspond to the mathematical theorems
- the proof procedure that is employed by Gorgias creates an admissible argument with references to specific rules and conditions that lead to the answer and this serves as an explanation towards the user
- The system is in the initial stages and can be extended with more methods or the methodology can be applied in different problems and other software packages

Thank You






-  Alexander M. Strasak, Qamruz Zaman, Karl P. Pfeiffer, Georg Göbel, Hanno Ulmer: Statistical errors in medical research—a review of common pitfalls. *Swiss Med Wkly*. 2007 Jan 27; 137(3-4): 44–49.
-  Murphy J.R.: Statistical errors in immunologic research, (2004) *Journal of Allergy and Clinical Immunology*, 114 (6) , pp. 1259-1263.
-  Young J., Statistical errors in medical research – a chronic disease?, *Swiss Med Wkly*. 2007;137:41–3.
-  Ilker Ercan, Berna Yazıcı, Yaning Yang, Guven Ozkaya, Sengul Cangur, Bulent Ediz, Ismet Kan (2007). Misusage of statistics in medical research. *European Journal of General Medicine* (ISSN: 1304-3897) Vol 4 Num 3. 4. [10.29333/ejgm/82507](https://doi.org/10.29333/ejgm/82507)



References ii

-  Hanif A, Ajmal T. Statistical errors in medical journals (A critical appraisal) Ann King Edward Med Univ. 2011;17:178–82.
-  Glenn T. Clark, Roseann Mulligan, Fifteen common mistakes encountered in clinical research, Journal of Prosthodontic Research, Volume 55, Issue 1, 2011, Pages 1-6, ISSN 1883-1958, <https://doi.org/10.1016/j.jpor.2010.09.002>.
<http://www.sciencedirect.com/science/article/pii/S188319581000112X>
-  Holmes TH1, Ten categories of statistical errors: a guide for research in endocrinology and metabolism. Am J Physiol Endocrinol Metab. 2004 Apr;286(4):E495-501.
<https://doi.org/10.1152/ajpendo.00484.2003>

-  A. C. Kakas, P. Mancarella, and Phan Minh Dung. 1994. The acceptability semantics for logic programs. In Proceedings of the eleventh international conference on Logic programming, Pascal Van Hentenryck (Ed.). MIT Press, Cambridge, MA, USA, 504-519.
-  Y. Dimopoulos, A. Kakas - Logic Programming without Negation as Failure, 1995, Proceedings of the Fifth International Logic Programming Symposium, ILPS-95, pp. 369-384, MIT Press, 1995.
-  28. A. Kakas, F. Toni, and P. Mancarella. Argumentation Logic. In Proceedings of the 5th International Conference on Computational Models of Argument (COMMA), pages 12-27, 2014.

References iv

-  N. I. Spanoudakis, E. Constantinou, A. Koumi and A. C. Kakas, "Modeling data access legislation with Gorgias," in 30th International Conference on Industrial, Engineering, and Other Applications of Applied Intelligent Systems, 2017, pp. 317-327. https://doi.org/10.1007/978-3-319-60045-1_34
-  A.C. Kakas, P. Mancarella; On the semantics of abstract argumentation, Journal of Logic and Computation, Volume 23, Issue 5, 1 October 2013, Pages 991-1015, <https://doi.org/10.1093/logcom/exs068>
-  Gorgias, An Argumentation System with Abduction, <http://www.cs.ucy.ac.cy/nkd/gorgias/>
-  Gorgias-B Argumentation Tool, <http://gorgiasb.tuc.gr/>
-  Fouskakis, D. Data Analysis using R, 2013, ISBN: 9786188074156

-  Fudenberg, Drew & K. Levine, David. (1996). The Theory of Learning in Games. David K. Levine, Levine's Working Paper Archive. 2.
-  Kiouvrekis, Y., Stefaneas, P. and A. Kokkinaki "An Argumentation-based Statistical Support Tool", In: Proceedings of the International Workshop on Applied Methods of Statistical Analysis, Nonparametric Methods in Cybernetics and System Analysis, Krasnoyarsk, The Russian Federation, September 18-22, 2017.