Title: Statistical Analyses for Single-Objective Stochastic Optimization Algorithms

Organizers: Tome Eftimov and Peter Korošec
Emails: tome.eftimov@ijs.si peter.korosec@ijs.si
Websites: http://cs.ijs.si/eftimov/, http://cs.ijs.si/korosec/
Affiliation: Computer Systems Department, Jožef Stefan Institute, Ljubljana, Slovenia

Brief abstract

Motivated by the great success of the previous tutorials entitled "Statistical Analyses for Meta-heuristic Stochastic Optimization Algorithms" held at GECCO 2020, 2021, 2022, IEEE CEC 2021, 2022 (WCCI 2022), and PPSN 2020, 2022, we hereby apply to organize a similar tutorial at IEEE CEC 2023. The scope of the IEEE CEC 2022 tutorial was on multi-objective optimization, while the scope of this proposal is on analysis for single-objective optimization.

Abstract

Moving to the era of explainable AI, a comprehensive comparison of the performance of single-objective stochastic optimization algorithms has become an increasingly important task. One of the most common ways to compare the performance of stochastic optimization algorithms is to apply statistical analyses. However, for performing them, there are still caveats that need to be addressed for acquiring relevant and valid conclusions. First of all, the selection of the performance measures should be done with great care since some measures can be correlated and their data is then further involved in statistical analyses. Further, the statistical analyses require good knowledge from the user to apply them properly, which is often lacking and leads to incorrect conclusions. Next, the standard approaches can be influenced by outliers (e.g., poor runs) or some statistically insignificant differences (solutions within some ε-neighborhood) that exist in the data.

This tutorial will provide an overview of the current approaches for analyzing algorithms' performance with special emphasis on caveats that are often overlooked. We will show how these can be easily avoided by applying simple principles that lead to Deep Statistical Comparison. The tutorial will not be based on equations, but mainly on examples through which a deeper understanding of statistics will be achieved. Examples will be based on various comparison scenarios for single-objective optimization algorithms. The tutorial will end with a demonstration of a web-service-based framework (i.e. DSCTool) for statistical comparison of single-objective stochastic optimization algorithms. In addition, R and Python clients for performing the analyses will be also presented.

Keywords
Robust statistics, benchmarking, single-objective optimization, metaheuristics

A tutorial description

To make the audience familiar with the terms that will be used during the tutorial we will begin by giving an introduction to statistical analysis. We will explain the difference between descriptive statistics, and inferential statistics (Frequentist vs. Bayesian). We will also provide the audience with a background in frequentist hypothesis testing, which is key to making a statistical comparison and always involves two hypotheses, the *null* and the *alternative*. Next, we will describe different types of statistical tests (e.g., parametric vs. non-parametric, omnibus vs. post-hoc) and discuss the required conditions that must be met in order to apply them properly. The selection of a statistical test is crucial to the outcome of a study because applying an inappropriate test can lead to a wrong conclusion. We will explicitly point out typical mistakes, which are often found in publications, and result from a lack of statistical knowledge. We will go into explaining the difference between the practical and the statistical significance and discuss how the performance measure (i.e., quality indicators) influences a statistical comparison with an emphasis on single-problem and multiple-problem analysis. This will be followed by a brief explanation of different statistical scenarios including pairwise comparison, multiple comparisons, and multiple comparisons with a control algorithm and we will follow this up by providing the audience with an overview of the standard approaches for making statistical comparisons and the latest advances (Deep Statistical Comparison) for providing robust statistical results. We will provide examples comparing single-objective stochastic optimization algorithms in different statistical scenarios. Finally, we will give an actual demonstration in which the audience will get to use a web-service-based framework for making a statistical comparison easier without having to worry about making incorrect conclusions. The tutorial will conclude with a summary of the covered topics and important take-home messages.

Outline of the covered material

1. Introduction to statistical analysis (Frequentist vs. Bayesian).
2. Background on frequentist hypothesis testing, different statistical tests, the required conditions for their usage, and sample size.
3. Typical mistakes and understanding why making a statistical comparison of data need to be done properly.
4. Understanding the difference between statistical and practical significance.
5. Understanding the effect that performance measures have on making a statistical comparison.
6. Defining single-problem and multiple-problem analysis.
7. Insight into pairwise comparison, multiple comparisons (all vs. all), and multiple comparisons with a control algorithm (one vs. all).
8. Standard approaches to making statistical comparisons and their deficiencies.
9. Latest advances in making statistical comparisons e.g., Deep Statistical Comparison, which provides more robust statistical results in cases of outliers and statistically insignificant differences between data values.
10. Extended Deep Statistical Comparison for understanding exploitation and exploration powers of stochastic optimization algorithms.
11. Examples of all possible statistical scenarios in single-objective optimization and caveats.

12. Presentation of a web-service-based framework that automatizes and simplifies the whole process of making a statistical comparison.
13. Take home messages.

Description of the tutorial demo

In the demo, we will present how the Deep Statistical Comparison web-service-based framework (i.e. DSCTool[1]) can be used in order to prevent us from drawing incorrect conclusions. The examples will include a comparison of single-objective optimization algorithms involved in different test scenarios: a pairwise comparison, multiple comparisons among all the algorithms, and multiple comparisons with a control algorithm. The demonstration will include examples using DSCTools web-service clients written in R and Python.

Classification of the tutorial

Advanced

Short bio of the instructors

Tome Eftimov is a senior researcher at the Computer Systems Department at the Jožef Stefan Institute. He is a visiting assistant professor at the Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Skopje. He was a postdoctoral research fellow at Stanford University, USA, where he investigated biomedical relations outcomes by using AI methods. In addition, he was a research associate at the University of California, San Francisco, investigating AI methods for rheumatology concepts extraction from electronic health records. He obtained his PhD in Information and Communication Technologies (2018). His research interests include statistical data analysis, metaheuristics, natural language processing, representation learning, and machine learning. He has been involved in courses on probability and statistics, and statistical data analysis. The work related to Deep Statistical Comparison was presented as a tutorial (i.e. IJCCI 2018, IEEE SSCI 2019, GECCO 2020, 2021, 2022, PPSN 2020, 2022, IEEE CEC 2021, 2022) or as an invited lecture to several international conferences and universities. He is an organizer of several workshops related to AI at high-ranked international conferences. He was a coordinator of a national project “Mr-BEC: Modern approaches for benchmarking in evolutionary computation”, currently a coordinator of a national project “Representation Learning of Landscape Spaces for Explainable Performance of Stochastic Optimization Algorithms”, and actively participates in European projects.

Peter Korošec received his Ph.D. degree from the Jožef Stefan Postgraduate School, Ljubljana, Slovenia, in 2006. Since 2002, he has been a researcher at the Computer Systems Department, Jožef Stefan Institute, Ljubljana. He developed a multilevel ant-colony optimization algorithm, which was tailored for mixed integer optimization problems with special emphasis on problems

[1] Eftimov, T., Petelin, G., & Korošec, P. (2020). DSCTool: A web-service-based framework for statistical comparison of stochastic optimization algorithms. *Applied Soft Computing*, *87*, 105977.

with multiple feasible regions and a small number of evaluations. He successfully applied his optimization approaches to several real-world problems in engineering. His current areas of research include understanding the principles behind meta-heuristic optimization. He participated in several tutorials related to statistical analysis for optimization algorithms presented at different international conferences (i.e. IJCCI 2018, IEEE SSCI 2019, GECCO 2020, 2021, 2022, PPSN 2020, 2022, IEEE CEC 2021, 2022) and co-organized a workshop on an understanding of evolutionary optimization behavior (IEEE CEC 2019) and automated algorithm design for evolutionary computation (IEEE CEC 2021, IEEE CEC 2022).

Potential target audience

The tutorial is planned for students and experienced researchers in the field of single-objective stochastic optimization algorithms. Many researchers have problems and difficulties in making a statistical analysis of their data, which they need to correctly interpret their results. To become familiar with making a proper statistical comparison, we propose a tutorial on how to perform a statistical comparison by focusing on state-of-the-art approaches that provide robust statistical results. We provide specific case studies where a statistical comparison is made using single-objective stochastic optimization algorithms. The proposed duration of the tutorial is 1.5 hours. The qualifications of instructors are described through a short bio and references presented at the end of the tutorial proposal.

Expected number of participants

Around 40

Details on previously held versions of the proposed tutorial including venues

The first appearance of the tutorial was at IJCCI 2018, which covers only the basic idea of Deep Statistical Comparison and the results that were published till 2018. The next appearances were at the IEEE SSCI 2019, GECCO 2020 (over 45 participants), and PPSN 2020 (over 40 participants) 2022 (over 60 participants), where the focus was on DSC analyses for both single- and multi-objective optimization algorithms. GECCO 2021 (over 40 participants) and IEEE CEC 2021 (over 35 participants) provided a modified version of the tutorial, which was on DSC analyses done for single-objective optimization. GECCO 2022 (over 30 participants) and IEEE CEC 2022 (over 30 participants) focused on DSC analyses in multi-objective optimization.

References:

Eftimov, T., & Korošec, P. (2022). Deep Statistical Comparison for Meta-heuristic Stochastic Optimization Algorithms", Springer Book

Eftimov, T., & Korošec, P. (2020, July). Statistical analyses for meta-heuristic stochastic optimization algorithms: GECCO 2020 tutorial. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion* (pp. 724-746).

Bartz-Beielstein, T., Doerr, C., Bossek, J., Chandrasekaran, S., Eftimov, T., Fischbach, A., ... & Naujoks, B. (2020). Benchmarking in optimization: Best practice and open issues. *arXiv preprint arXiv:2007.03488*.

Korošec, P., & Eftimov, T. (2020). Multi-Objective Optimization Benchmarking Using DSCTool. *Mathematics*, *8*(5), 839.

Eftimov, T., Petelin, G., Hribar, R., Popovski, G., Škvorc, U., & Korošec, P. (2020, July). Deep statistics: more robust performance statistics for single-objective optimization benchmarking. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion* (pp. 5-6).

Eftimov, T., Popovski, G., Kocev, D., & Korošec, P. (2020, July). Performance2vec: a step further in explainable stochastic optimization algorithm performance. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion* (pp. 193-194).

Eftimov, T., Hribar, R., Škvorc, U., Popovski, G., Petelin, G., & Korošec, P. (2020, July). PerformViz: a machine learning approach to visualize and understand the performance of single-objective optimization algorithms. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion* (pp. 7-8).

Korošec, P., & Eftimov, T. (2020). Insights into Exploration and Exploitation Power of Optimization Algorithm Using DSCTool. *Mathematics*, *8*(9), 1474.

Eftimov, T., & Korošec, P. (2020, July). Statistical analyses for meta-heuristic stochastic optimization algorithms: GECCO 2020 tutorial. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion* (pp. 724-746).

Eftimov, T., & Korošec, P. (2019). Identifying practical significance through statistical comparison of meta-heuristic stochastic optimization algorithms. *Applied Soft Computing*, 105862.

Eftimov, T., & Korošec, P. (2019). A novel statistical approach for comparing meta-heuristic stochastic optimization algorithms according to the distribution of solutions in the search space. *Information Sciences*, *489*, 255-273.

Eftimov, T., & Korošec, P. (2019). Understanding exploration and exploitation powers of meta-heuristic stochastic optimization algorithms through statistical analysis. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (pp. 21-22). ACM.

Škvorc, U., Eftimov, T., & Korošec, P. (2019). GECCO black-box optimization competitions: progress from 2009 to 2018. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (pp. 275-276). ACM.

Škvorc, U., Eftimov, T., & Korošec, P. (2019, June). CEC Real-Parameter Optimization Competitions: Progress from 2013 to 2018. In *2019 IEEE Congress on Evolutionary Computation (CEC)* (pp. 3126-3133). IEEE.

Eftimov, T., Korošec, P., & Seljak, B. K. (2018). Deep Statistical Comparison of Meta-heuristic Stochastic Optimization Algorithms. In *Proc. Companion, Genetic and Evolutionary Computation Conference*, Kyoto, Japan.

Eftimov, T., & Korošec, P. (2018). The Impact of Statistics for Benchmarking in Evolutionary Computation Research. In *Proc. Companion, Genetic and Evolutionary Computation Conference*, Kyoto, Japan.

Eftimov, T., & Korošec, P. (2018). Tutorial: How to Perform a Proper Statistical Analysis? Where we Started and Where are we Now in statistical Performance Assessment Approaches for Stochastic Optimization Algorithms? At *10th International Conference on Computational Intelligence*, Sevillie, Spain.

Eftimov, T., Korošec, P., & Seljak, B. K. (2018). Data-Driven Preference-Based Deep Statistical Ranking for Comparing Multi-objective Optimization Algorithms. In *International Conference on Bioinspired Methods and Their Applications* (pp. 138-150). Springer, Cham.

Eftimov, T., Korošec, P., & Seljak, B. K. (2017). A Novel Approach to statistical comparison of meta-heuristic stochastic optimization algorithms using deep statistics. *Information Sciences*, *417*, 186-215.

Eftimov, T., Korošec, P., & Seljak, B. K. (2017). Comparing multi-objective algorithms using an ensemble of quality indicators with deep statistical comparison approach. In the *2017 IEEE Symposium Series on Computational Intelligence*, Honolulu, Hawaii.

Eftimov, T., Korošec, P., & Seljak, B. K. (2017). Deep Statistical Comparison applied on quality indicators to compare multi-objective stochastic optimization algorithms. In Proceedings of the *3th International Conference on Machine Learning, Optimization, and Big Data*, Volterra, Italy.

Eftimov, T., Korošec, P., & Seljak, B. K. (2017). The Behavior of deep statistical comparison approach for different criteria of comparing distributions. In Proceedings of the *9th International Conference on Computational Intelligence*, Madeira, Portugal.

Eftimov, T., Korošec, P., & Seljak, B. K. (2016). Disadvantages of Statistical Comparison of Stochastic Optimization Algorithms. *Proceedings of the Bioinspired Optimization Methods and their Applications, BIOMA*.

Web site for DSC tool: https://ws.ijs.si:8443/dsc-1.5/documentation.pdf