

Practical application of the Wilcoxon-Mann-Whitney test in valuation.

**Selection of attributes as pricing factors based on the principle of unbiased
estimates**

K. A. Murashev

July 22, 2022

In their practice appraisers often face the need to take into account differences in quantitative and qualitative characteristics of objects. In particular, one of the standard tasks is to determine the attributes that influence the cost (so-called "pricing factors") and to separate them from the attributes that do not or cannot be determined.

Subjective selection of attributes taken into account in determining the value is widespread in valuation practice. In this case, specific quantitative indicators of the impact of these attributes on the cost are often taken from the so-called "reference books". While not denying the speed and low cost of this approach, it should be recognized that only data directly observed in the open markets is a reliable basis for a value judgment. The priority of such data over other data, in particular those obtained by expert survey, is enshrined, among others, in RICS Valuation — Global Standards 2022 [**RVGS-2022**], International Valuation Standards 2022 [**IVS-2022**], as well as in IFRS 13 "Fair Value Measurement" [**IFRS-13**]. Therefore, we can say that mathematical methods for analyzing data from the open market are the most reliable means of interpreting market information used in market research and predicting the value of individual objects.

The aim of this work is to justify the necessity and possibility of using a rigorous mathematical Wilcoxon–Mann–Whitney test, which allows us to answer the question about the necessity of taking into account the binary attribute as a price-generating factor. Instead of the judgmental approach, which is most commonly used by appraisers in selecting the attributes to be considered in appraisal, this paper proposes the idea of prioritizing the measuring approach based on the results of a mathematical test that allows to draw a conclusion about the importance or otherwise of the binary attribute influence on the value. It should be noted that despite the fact that the statistical test under consideration belongs to frequentist statistics, it, through its connection to ROC analysis and AUC, is related to modern machine learning methods, which will be discussed later in the text of this material. The presence of this relationship and elements of Bayesian statistics seems particularly interesting and promising from the point of view of introducing machine learning and data analysis methods into the everyday practice of appraisers.

Users should have some general math background and basic Python and R programming skills to understand and practice all of the material in the text, but lack of that knowledge and skill is not a barrier to learning most of the material and implementing the test in the spreadsheet that comes with it.

The material consists of four blocks:

- a description of the Wilcoxon–Mann–Whitney test (hereafter "U-test"), its probabilistic meaning, and its relationship to other mathematical methods;
- a practical implementation of the U-test in a spreadsheet on an example of test random data;

- practical implementation of the U-test on the real data of the residential real estate market of St. Petersburg agglomeration by means of Python programming language, the purpose of the analysis was to check the significance of the difference in the unit price between the objects located in the urban and suburban parts of the agglomeration;
- practical implementation of the U-test on real data of residential real estate market of Almaty by means of R programming language, the purpose of the analysis was to check the significance of difference in unit price between the objects sold without demountable improvements and the objects sold with them.

The current version of this material, its source code, Python and R scripts, and the spreadsheet are in the repository on the GitHub portal and are available at the permanent link [**Murashev:u-test**].

This material and all of its appendices are distributed under the terms of the cc-by-sa-4.0 license [**cc-by-sa-4.0**].

Contents

List of Tables

List of Figures

Listings