SAT-9.3-1-HP-16

Analysis of Quantitative Data Obtained by Questionnaire Studies

Nushka Kehayova, Detelin Markov

Abstract: *This paper deals with analysis of quantitative data obtained from human subject experiments by questionnaires. Typically the analysis in such situation is related to testing the distribution of the studied variables and/or identifying the differences between two or more of them. It is very important to determine if the observed differences between investigated variables are result of the influence of random factors and events or if there is a statistically significant relation between them. The current paper presents an example for statistical analysis of quantitative data, obtained from a questionnaire study. The subjective vote of the participants in the study is investigated for statistically significant difference using several tests. Based on the results from the performed statistical analysis different assumptions (hypothesis) regarding the subjective vote are tested.*

Key words*: Human Subject Experiment, Qualitative Data, Statistical Analysis, Non-Parametric Tests*

INTRODUCTION

In human subject experiments related to studying the influence of the indoor environment on the health, comfort and productivity of the tested subjects are collected and analyzed various types of data: objectively measured and subjectively evaluated variables characterizing the physical environment as well as objectively measured and subjectively evaluated variables characterizing the subjects - both physiological and psychological. In such experiments it is always important to understand how the test subjects rate and perceive the elements of the indoor environment – thermal environment, indoor air quality, etc.

Under these experiments human subjects evaluate, “measure”, the parameters of the indoor environment by a set of Likert scales (Rensis Likert, 1932). An example of a Likert scale is the EN 15251 (2007) standard, [1], acceptability scale, presented on Figure 1. Though there are only 4 answers marked on this scale it is proportional in nature and the answers obtained by it are values of a quantitative variable, [5].

There is no zero (a neutral point) on this scale. It is composed of two wings, acceptable and unacceptable, with equal length. This scale forces the test subjects firstly to take a decision about the acceptability of the investigated parameter of the indoor environment and only then to measure it by putting a mark on the line in the selected wing. In order to process and interpret the results, the answers obtained by this scale, are coded with numbers. For the scale presented on Figure 1 the values of the boundary points are as follows:*”Just acceptable” = 0.01, ”Clearly acceptable” = 1, ”Just unacceptable” = -0.01, and ”Clearly unacceptable” = -0.01.* The ratio of the distance between the mark and the beginning point of the selected wing of the scale to the length of the wing is proportional to the level of acceptability/unacceptability of the studied parameter.

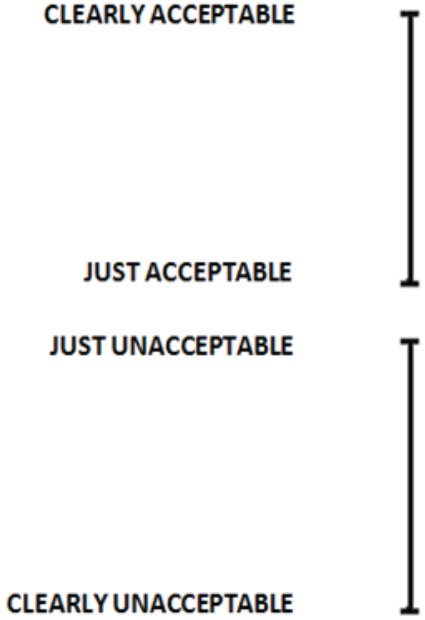

Figure 1: EN 15251 (2007) acceptability scale.

The goal of this paper is to present an example for statistical analysis of quantitative data, obtained by such scales under a human subject experiment organized to evaluate the environment in two types

of hospital beds.

METHODS

The response of 32 subjects (17 males and 15 females) to the local environment generated at a standard hospital bed and a Hospital Bed with Installed Ventilation and air Cleaning Unit (HBIVCU) was studied, [3]. The HBIVCU is an advanced air distribution system that supplies clean air close to the patient's breathing zone and exhausts the polluted (might be infected) air from the patient's pulmonary activities (breathing, coughing), thus reducing the risk of airborne cross-infection for all room occupants (patients, doctor, nurses, etc.), [4, 5]. Two hospital beds, one with (VB-Ventilated Bed) and one without a HBIVCU (NVB-Non-Ventilated Bed), were located in two separate rooms that mimic a single person hospital room with dimensions of 6 m x 3 m x 3 m (L x W x H). The thermal conditions in both rooms were identical: air temperature was 23±1°C and relative humidity 42±1.5%. The rooms were ventilated by mixing type ventilation at 9 ACH (Air Changes per Hour). Each tested subject spent two hours in two different days in each of the rooms. The duration of the experiment was divided into three periods. The first period, 30 minutes long, was designed to ensure acclimatization of the test subjects to the environment in the room but not in the bed. During the second period, 60 minutes long, test subjects were exposed to the environment in the bed. During this period, the response of each subject to the environment in each bed was collected four times (at t_{1t}=30 min, t_{2t}=50 min, t_{3t}=70 min, and t_{4t}=90 min) via questionnaires. The subjects were asked to evaluate the air quality and the thermal environment in the bed by means of the acceptability scales presented in EN 15251 (2007) standard, [1].

The study was designed to find if there is a difference between the environment in the VB and NVB based on the perception of the test subjects. In order to answer this question the analysis could be composed of the following steps (giving an answer to each of the following questions):

Q1: What is the distribution of the evaluations of the test subjects in the two beds at each moment of time?
Q2: Is there a (statistically significant) difference between the evaluations of the group of test subjects as a whole for the environment in the VB and the NVB?
Q3: Is there a (statistically significant) time variation of the evaluations of the group of test subjects as a whole for the environment in the VB and the NVB?
Q4: Is there a (statistically significant) difference between the votes of male and female subjects in the VB and in the NVB?

DATA ANALYSIS AND RESULTS

Figure 2 shows the spread of the individual subjective vote over the exposure period in the VB and the NVB with respect to the air quality acceptability and thermal sensation acceptability. Based on the figure, the environment in the two beds at the four moments of the exposure of each subject may be visually compared, but no significant conclusions regarding the differences in the subjective vote can be made. In order to answer the questions listed in the Methods section the data have to be analyzed statistically. Prior to applying any test for statistical significance, the distribution (Gaussian or Non-Gaussian) of each data sample of 32 values (collected from 32 subjects) has to be identified. Here is analysed only the Perceived air quality acceptance vote of the test subjects.

Q1: What is the distribution of the evaluations of the test subjects in the two beds at each moment of time?

To determine the distribution type of the evaluations in this study the Shapiro-Wilk test is used. Two hypotheses are stated: the null hypothesis H_0, which claims that the data are normally distributed and the alternative hypothesis H_1, which assumes that the data

are not normally distributed. The level of significance is accepted to be α=0.05. If the result of the test is a p-value less than α, i.e. p<0.05, the null hypothesis is rejected and the tested data sample is considered to be not-normally distributed. The Shapiro –Wilk test applied to all 8 sets of evaluations, presented on Fig.2, shows that they are not-normally distributed. Hence, for answering the other 3 questions (Q2 – Q4) only non-parametric tests have to be used.

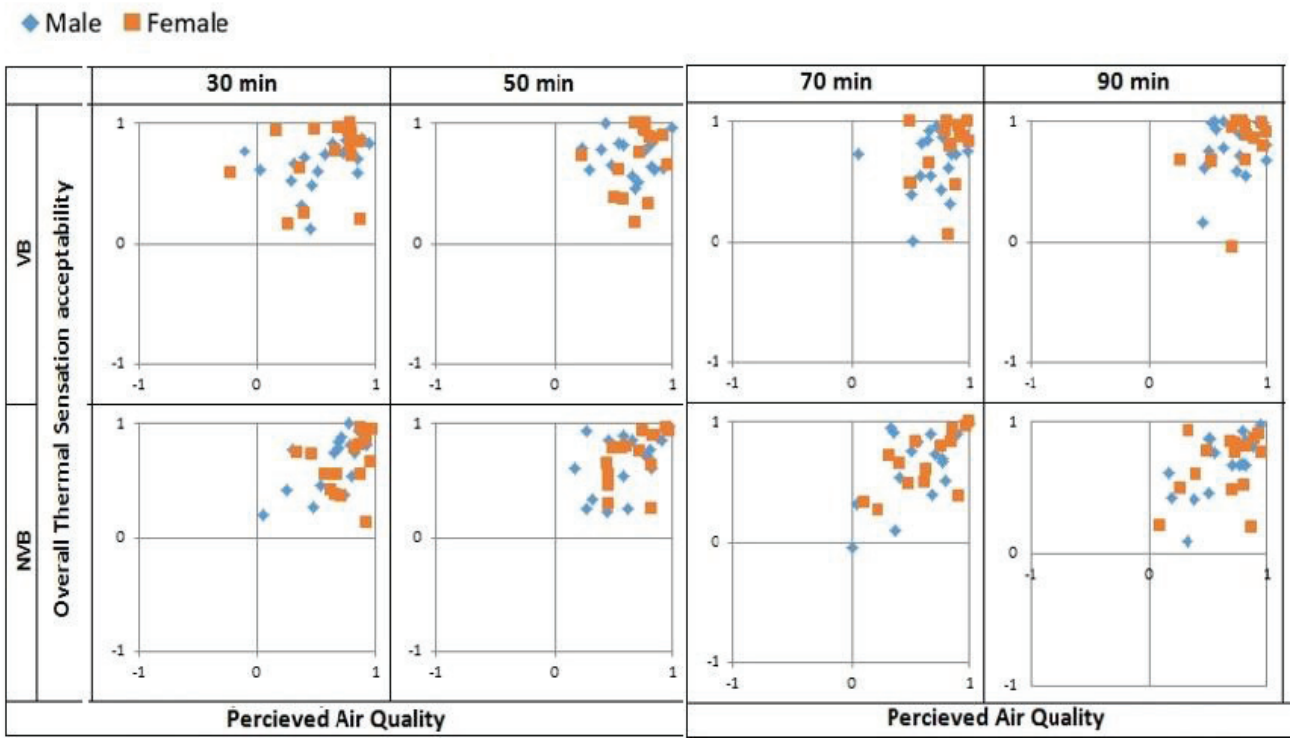

Figure 2: Individual subjective vote time change over the exposure period

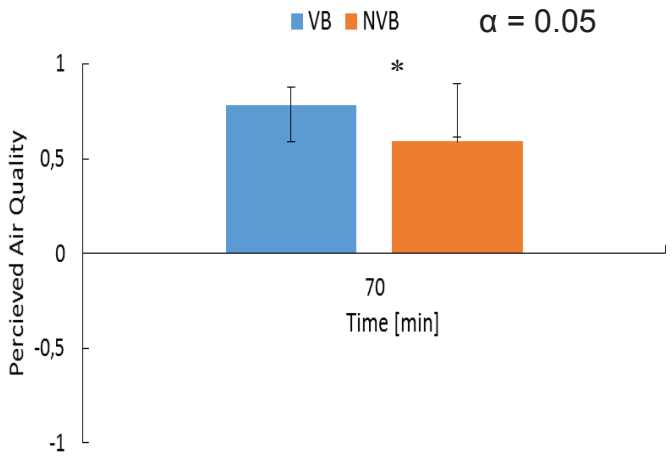

Figure 3: Median Perceived Air Quality Acceptability vote for the 70-th min of the exposure period

Q2: Is there a (statistically significant) difference between the evaluations of the group of test subjects as a whole for the environment in the VB and the NVB?

To answer that question the Wilcoxon Signed Rank Test has to be applied, [2, 7, 8]. It is a nonparametric test for comparing dependent variables. Wilcoxon Signed Rank Test is an alternative for the t-test in the case of normally distributed data. Evaluations of the test subjects at each moment (t_{1t}, t_{2t}, t_{3t}, and t_{4t}) for the environment in the VB are compared with their evaluations for the environment in the NVB. Null hypothesis for this test is that there is no significant difference. The level of significance is α=0.05. The results from the test are as follows: at t_{1t} p=0.040, at t_{2t} p=0.059, at t_{3t} p=0.010, and at t_{4t} p=0.031. This means that at t_{1t}, t_{3t}, and t_{4t} there is a statistically significant difference between the evaluations of the test subjects for the environment in the VB and NVB, while at t_{2t} this difference is not statistically significant.

On Figure 3 is presented the median of the evaluations of the test subjects for the environment in the VB and NVB at moment t_{3t} together with the 25-th and 75-th percentile. The Wilcoxon signed rank test for this moment shows that the Perceived Air Quality Acceptability vote in the VB case is significantly higher than the one for the NVB with p=0.010.

Q3: Is there a (statistically significant) time variation of the evaluations of the group of test subjects as a whole for the environment in the VB and the NVB?

To compare three of more dependent samples in a case of not normally distributed data the Friedman (ANOVA) test can be used, [2, 7, 8]. The test gives ranks to all variables for the different trials and is used to evaluate hypothesis over a period of time. In this study the test was used to determine if there is a statistically significant time variation of the evaluations of the group of test subjects as a whole for the environment in the VB as well as in the NVB. Null hypothesis (H_0) of the test is that there is no time variation of the evaluations for the environment in the considered bed. The level of significance is accepted to be α=0.05. If p<α then the alternative hypothesis (H_1) of differences between the compared trials ıs accepted.

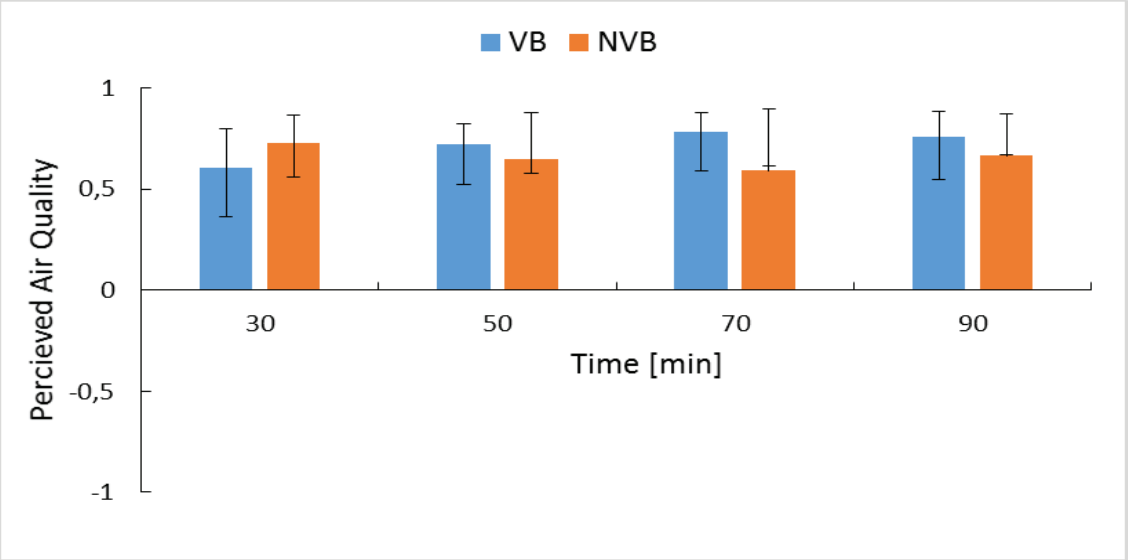

Fig.4 Median Perceived Air Quality Acceptability vote time variation over the exposure period

Figure 4 presents the median Perceived Air Quality Acceptability vote at the preset moments of the exposure (t_{1t}=30 min, t_{2t}=50 min, t_{3t}=70 min, and t_{4t}=90 min) for the VB and the NVB case. The error bars on the figure show the 25-th and 75-th percentiles. Applying the Friedman test for the VB case showed that the Perceived Air Quality Acceptability vote had a tendency to increase over time with p-value p=0.000. Based on this result the alternative hypothesis H_1 for significant difference in the subjective vote over time was accepted. However, applying the Friedman test to the NVB resulted in p=0.719. Hence, null hypothesis (H_0) of the test that there is no time variation of the evaluations for the environment in the considered bed is accepted.

Q4: Is there a (statistically significant) difference between the votes of male and female subjects in the VB and in the NVB?

In order to identify differences between the subjective vote of the male and the female subjects the Mann-Whitney U test has to be applied, [2, 8]. The Mann-Whitney U test is the non-parametric alternative to the independent sample t-test. Mann-Whitney U

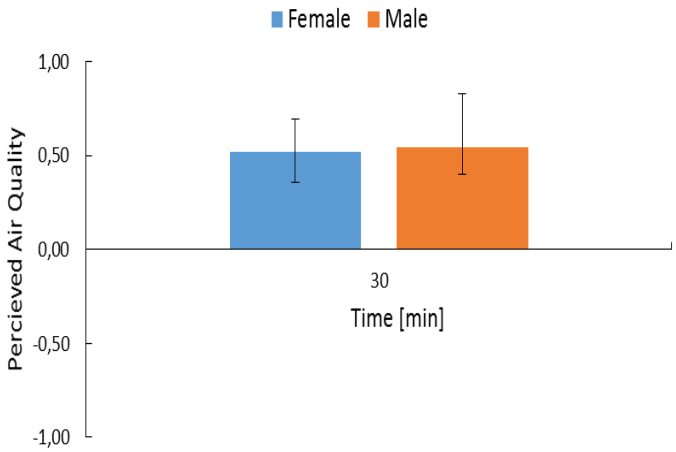

Fig. 5 Median Perceived Air Quality Vote of the male and female subjects for the 30-th min of the exposure period

test is used to compare independent variables, [2, 8]. It makes no assumptions related to the distribution. However, the following considerations are taken: 1.) the sample drawn from the population is random; 2.) there is mutual independence and independence within the samples; 3.) at least ordinal measurement scale must be used for collection of the data (i.e. one can say, out of any two observations, which one is greater). Null hypothesis of the test (H_0) is that there is no significant difference between the votes of males and females. The level of significance was accepted to be α=0.05. The evaluations of the male and female test subjects for the environment in each bed at each preset moment of the exposure interval meet the requirements of the Mann-Whitney U test.

Figure 5 presents comparison between the median Perceived Air Quality Acceptability voted by the male and the female subjects at the 30-th min of the exposure period. The error bars shown on the figure are the 25 and 75 percentiles. The results from the applied statistical test showed that the level of significance was p=0.955. Since the p-value is greater than 0.05 the null hypothesis H_0 is proved – there is no significant difference in the subjective votes of the male and female participants in the experiment in the VB case at the 30-th min of the exposure period. For all other cases the result is the same. The smallest p-value p=0.719 is obtained for the NVB at the moment t_{3t}.

CONCLUSION

This paper presents an example for statistical analysis of quantitative data obtained under a human subject experiment. Presented examples reveal that even when the subjective votes are grouped in a similar way (Figure 2) there could exists a statistically significant difference between the group votes. They reveal as well that even when the data are visually different (Figure 3) not always this difference is statistically significant.

References

[1] DS/EN 15251, (2007) “Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lightning and acoustics”, 1-st edition, European Committee for Standardization

[2] Siegel S., N.J. Castellan Jr, Nonparametric statistics for the behavioral sciences, McGraw-Hill, 1988, ISBN 0-07-057357-3

[3] Kehayova N, Bolashikov Z., Melikov A., “ Subjective Evaluation of the Microenvironment Generated by a Hospital Bed with Localized Ventilation System”, Indoor Air Conference, Ghent, Belgium, , July 2016

[4] Melikov, A., Bolashikov, Z., Brand, M., (2010). “Experimental investigation of performance of a novel ventilation method for hospital patient rooms.” 21st Congress of International Federation of Hospital Engineering (IFHE), Tokyo Japan, November 17th to 19th, 2010

[5] Melikov, A., Bolashikov, Z., Georgiev E., (2011). “Novel ventilation strategy for

reducing the risk of cross infection in hospital rooms." In: Proceedings of Indoor Air 2011.Paper 1037.
[6] Questionnaire design and data analysis using SPSS, SPSS Inc. 2004
[7] Analysis of Questionnaires and Qualitative Data, Non- Parametric Tests, Jerzy Stefanowski, Instytut Informatyki Politechnika Poznańska, Lecture notes (2013)
[8] Kenny D.A., Statistics for the social and behavioral sciences, ISBN/ASIN: 0316489158, Little, Brown 1987.

Contact information:

Nushka KEHAYOVA, MSc, **PhD Student,** Department of Hydroaerodynamics and Hydraulic Machines, Technical University of Sofia, Bulgaria, e-mail: n.kehayova@abv.bg

Assoc. Prof. Detelin MARKOV, MSc, **PhD,** Department of Hydrodynamics and Hydraulic Machines, Technical University of Sofia, Bulgaria, e-mail: detmar@tu-sofia.bg

This paper has been reviewed.