Websites for Hypothesis Testing

František Mošna

Abstract—E-learning has become an efficient and widespread means of education at all levels of human activities. Statistics is no exception. Unfortunately the main focus in statistics teaching is usually paid to the substitution in formulas. Suitable websites can simplify and automate calculations and provide more attention and time to the basic principles of statistics, mathematization of real-life situations and following interpretation of results. We now introduce our own web-site for hypothesis testing. Its didactic aspects, the technical possibilities of the individual tools, the experience of use and the advantages or disadvantages are discussed in this paper. This web-site is not a substitute for common statistical software but should significantly improve the teaching of statistics at universities.

Keywords—E-learning, hypothesis testing, PHP, websites.

I. INTRODUCTION

SINCE the end of past century new technologies for teaching using computers and the web has appeared in the process of education at all types and levels of schools in all countries (see e.g. [5], [8]). There are plenty of websites which offer text in book format but they hardly differ from actual printed books. However, some advanced hypertexts (e.g. multiple choice tests, animated or interactive figures etc.) do make use of the significant advantage of internet pages – the interactivity. We would like to introduce a web-site for statistical method called hypothesis testing. The contents of this site are accessible at the web address http://statisticsonweb.tf.czu.cz.

II. WEBSITES FOR STATISTICS TEACHING

Statistics seems to be one of the less popular branches of mathematics. The teaching of this discipline at basic and secondary schools is usually reduced only to the calculation of the mean or the standard deviation. ”Real” mathematical statistics is first taught at university. Statistics is a discipline which gains, collects, processes and clearly presents the bulk data and it tries to acquire some important results and information from them. So statistics is very useful in spite of not being very popular.

From long experience we are aware that the greatest effort is focused on calculations according to the given formulas (for the means, the standard deviations etc.) rather than on the basic ideas of method or evaluation, mathematization and interpretation of real situations. Hence we were looking for some tools which could simplify and automate statistical calculation and bring more attention and time to the core of hypothesis testing during lessons.

There are some utilities in MS Excel or MATHEMATICA and plenty of special statistical software e.g. STATISTICA, SPSS, SAS, Statgraphic, R etc. But to use such programs students need to acquire certain additional skills in how to handle them, which can again divert attention from the core of statistical topics. Even on the web (with some exceptions) we did not find appropriate tools which could enable students to automate the calculation part of problem solving at school or during their off-class learning.

That is why we decided to create a special web-site which could simplify the process of calculations and which is easily manageable at the same time. The aim of this web-site is to emphasize the process of hypothesis testing, mathematization of the real situation, following interpretation and to totally simplify the technical parts of problem solving. We do not want to bother the student with technically demanding calculations or knowledge needed for manipulation with some statistical program.

III. TWO IMPORTANT MOMENTS IN MATHEMATICAL STATISTICS

A. Hypothesis Testing

One of the main procedures of mathematical statistics is hypothesis testing. Let us remember the basic principle of hypothesis testing (see e.g. [1]-[4] or [6]). We consider a certain initial hypothesis called null hypothesis. Our task is to decide on the basis of collected data whether this hypothesis can be rejected or not. It is important to realize that not rejecting does not mean that we accept the hypothesis. A similar situation occurs in law. The presumption of innocence of the accused is similar to the null hypothesis. If we reject the innocence of the accused he will be punished. If not, it does not mean he is really innocent. Either he is innocent or we have not enough evidence for his conviction. It is the same with hypothesis testing. One possibility is to reject the null hypothesis and we state that it is not true. Another possibility is not to reject it which can mean either it is true or we have not enough measurement to reject it.

This is the basic principle of the method of hypothesis testing which should be well understood by university students.

B. Mathematization and Interpretation

The second very important task during the teaching of statistics is to make students acquainted with capabilities which can enable them to orient themselves in real situations. Students should be able to mathematize real situations, to choose the right test and formulate correct null hypothesis and to be able to interpret the results.

F. Mošna is with the Department of Mathematics, Czech University of Life Sciences, Praha, the Czech Republic (e-mail: mosna@tf.czu.cz).

IV. Layout of the Website

Firstly, we had to consider the main features of the required web-site. It should consist of three parts: 1. selection of the input range, 2. entering the input data and 3. calculation. The main idea of a web-site is as follows. The user can enter range of random sample at first. Then the required number of items is opened and the user can enter data. After submitting data, the calculation is performed and the result is stated. For instance, at the one-sample test the user can select the range of random sample, i.e. the number of measurements available. The examples for teaching statistics should not be very large. Hence we permit the range from 2 to 30 cases. Another reason for the small range is that the Student's t-test (and others) loses its sense for a great number of measures because the Student's distribution approach to normal distribution according to the Central Limit Theorem. After submitting the range, the corresponding number of items is opened and it is possible to enter data to a form. Also, we can enter a hypothetical value with which we compare the mean. Additional submitting starts the calculation. The web-site displays values of the necessary statistics, the critical values (or quantiles) of the Student's distribution, the decision about rejecting or not rejecting the null hypothesis and the confidence interval. Both the range and the data can be changed and a new entry can be selected, too.
The website contains the following tests:

A. Tests Comparing Mean or Variance of One Sample with Some Given Value

- One-sample Student t-test two-tailed
- One-sample Student t-test one-tailed
- One-sample (Pearson) test for variance
- One-sample Wilcoxon test

B. Tests Comparing Means or Variances of Two Samples

- Two-sample Student t-test
- Paired t-test
- Two-sample Fischer test for variance
- Two-sample Wilcoxon (Mann-Whitney) test

C. Tests Comparing Means of More Samples

- One-way analysis of variance (ANOVA)
- Kruskal test
- Two-way analysis of variance (ANOVA) without interactions
- Friedman test
- Two-way analysis of variance (ANOVA) and interactions

D. Independence Tests

- Pearson correlation coefficient test
- Spearmann correlation coefficient test
- Contingency tables (crosstab)

The websites can be enriched by some more tests, for instance Wald-Wolfowitz test, Kendal correlation coefficient test, tests for partial correlation coefficients, McNemar test and especially linear regression, which demands to use graphical methods of PHP.

Any test is accompanied by file with used formulas. Hence the users can prove the results by own calculation.

V. Technical Aspects of the Website

Another equally important task was to decide which tools we should use for creating such a web-site. There are several technical possibilities to animate websites under the protocol HTML; the most common ones are the commands from the scripting languages JavaScript and PHP or a combination of both. The main difference between these two tools is that the first one works on the side of a server while the second one works on the side of a client. Both can be implemented with the HTML code. We found the PHP is better for this purpose and therefore we chose this tool for our work. We are going to also study the possibilities and advantages of JavaScript for the same task.

During our work on this web-site, it was necessary to focus on the solution of several part tasks. There are some of them:

A. Step by Step Opening Websites

This effect can be achieved simply and naturally in PowerPoint (at Windows) or Beamer (at LaTeX). In the past we used Java Script commands to gradually open examples in one of my earlier composed websites for e-learning in a basic course of maths. PHP offers switches for this purpose which can, together with elements <input type=hidden> from forms, enable the required effect. It can be implemented by the following commands:

```
<?php switch($a):
case 0: ?>   submitting of range with the form containing
<input type=hidden name=a value=1>      <?php
break;
case 1: ?> submitting of range and data with the form
containing <input type=hidden name=a value=2>
<?php break;
case 2: ?>      submitting of range, data and calculation with
the form containing <input type=hidden name=a
value=0>      <?php default:
endswitch;?>
```

B. Simulation of Two-Dimensional Array in HTML

Contrary to PHP, the forms in HTML do not use two dimensional arrays. Hence it was necessary to load data to one-dimensional array and then convert them into the multi-dimensional arrays of PHP.

C. Cooperating with Data Files

We found some possibility of calculating critical values (or quantiles) of distributions. PHP very effectively interacts with both text files and the databases MySQL. But we did not know how to generate the critical values of the necessary distributions (normal, Pearson's chi-square, Student's and Fischer's). We did not find the required statistical functions in neither PHP nor databases manager MySQL (which interacts with PHP). Maybe there are some suitable libraries of statistical functions which are supported by PHP. In the end we used the simple cooperation of PHP with ordinary text files and the web-site obtained the critical values from a type of

statistical table. It seems to be the simplest way to get these values. However, consequently the level of the test cannot be arbitrary but only 0.05.

VI. EXAMPLES

We will introduce the website and its advantages on two examples.

A. Example of One-Sample One-Tailed Student t- Test

The concentration of an acid solution must not exceed 20%. We have nine control measures:

20.3 20.4 19.8 20.5 19.7 20.3 20.5 19.9 20.4   (%).

Is the concentration of the solution systematically exceeded? (We choose standardly the test level 0.05.)

To decide and answer the question, we will assume that the data consist of random sample from normal distribution $N(\mu;\sigma^2)$ and we use one-sample Student t-test. We reject the null hypothesis $H_0: \mu \le 20$ because $T \ge t_8(0.95)$, where

$$T = \frac{\bar{X} - 20}{\sqrt{S_X^2/n}} = 1.922 , \quad \bar{X} = \frac{1}{n}(X_1 + \cdots + X_n) = 20.2 ,$$

$$S_X^2 = \frac{1}{n-1}((X_1 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2) = 0.0975$$

and $t_8(0.95) = 1.833$ is quantile of Student distribution. So we can state that the concentration of solution is exceeded.

Now we compare the calculation using various programs and calculation on the website (see e.g. [7]).

The simplest procedure in Excel is perhaps to calculate necessary variables using statistical functions.

→ T-Test

[DataSet0]

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
concentration	9	20,2000	,31225	,10408

One-Sample Test

	Test Value = 20					
				Mean Difference	95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)		Lower	Upper
concentration	1,922	8	,091	,20000	-,0400	,4400

Fig. 1 Example A (one-sample t-test) in SPSS

	Test of means against reference constant (value) (Spreadsheet1)							
Variable	Mean	Std.Dv.	N	Std.Err.	Reference Constant	t-value	df	p
Var3	20,20000	0,312250	9	0,104083	20,00000	1,921538	8	0,090900

Fig. 2 Example A (one-sample t-test) in STATISTICA

The results in SPSS and STATISTICA are obvious (see Figs. 1 and 2).

Now we shall use the website. When we enter range (=9) then nine windows is opened and we can enter values of sample, one window for required value of concentration (=20) and we can choose type of inequality, too. After we click button "perform the test" all necessary variables are calculated, corresponding decision is made and moreover interval of confidence is displayed (see Fig. 3).
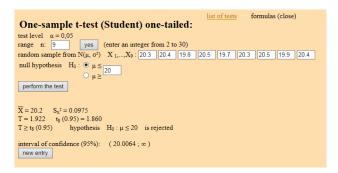

One-sample t-test (Student) one-tailed: list of tests   formulas (close)

test level $\alpha = 0{,}05$
range   n: [9]   [yes]   (enter an integer from 2 to 30)
random sample from $N(\mu, \sigma^2)$ $X_1, ..., X_9$ [20.3] [20.4] [19.8] [20.5] [19.7] [20.3] [20.5] [19.9] [20.4]
null hypothesis H_0 : ⦿ $\mu \le$ [20]
⦿ $\mu \ge$

[perform the test]

$\bar{X} = 20.2$ $S_e^2 = 0.0975$
$T = 1.922$ $t_8(0.95) = 1.860$
$T \ge t_8(0.95)$ hypothesis $H_0 : \mu \le 20$ is rejected

interval of confidence (95%):   (20.0064 ; ∞)
[new entry]

Fig. 3 Example A (one-sample t-test) on the website

B. Example of One-Way ANOVA

We would like compare the fuel consumption of three types of cars Fiat, Citroen and Renault. We have an average consumption of seven Fiats, seven Citroens and five Renaults:

Fiat   7.0 6.5 7.2 6.7 6.6 7.0 7.3
Citroen   7.1 7.2 6.9 7.0 7.3 7.2 7.0
Renault   7.5 7.2 7.4 7.6 7.3   (l/100 km).

Is the petrol consumption of the three types of cars equal? (We choose again the test level 0.05.)

To decide and answer the question, we will assume that the data consist of random samples from normal distributions $N(\mu_1;\sigma^2)$, $N(\mu_2;\sigma^2)$ and $N(\mu_3;\sigma^2)$, we suppose equality of the variances, and we use one-way ANOVA test. We reject the null hypothesis $H_0: \mu_1 = \mu_2 = \mu_3$, because $F \ge F_8(0.95)$, where

$$F = \frac{S_A/f_A}{S_e/f_e} = 1.922 ,$$

S_A and S_e are so called sums of squares, f_A and f_e are so called degrees of freedom and $F_8(0.95) = 1.833$ is quantile of Fischer distribution. So we can state that the petrol consumptions of the three types of cars differ. So called Sheffé post-hoc test yields us that difference between Fiat and Renault is significant.

Now we again compare the calculation using various programs and calculation on the website.

Excel contains suitable procedure for ANOVA (see Fig. 4)

The results in SPSS and STATISTICA are obvious (see Figs. 5 and 6).

Now we shall use the website. At first we enter number of classes (=3), then we enter range of the classes (=7,7,5). Corresponding number of windows is opened and we can enter values of sample. After we click button "perform the

test" all necessary variables (post-hoc test included) are calculated, displayed and corresponding decision is made (see Fig. 7).

Anova: Single Factor				
SUMMARY				
Groups	Count	Sum	Average	Variance
Fiat	7	48,3	6,9	0,093333
Citroen	7	49,7	7,1	0,02
Renault	5	37	7,4	0,025
ANOVA				

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0,729474	2	0,364737	7,481781	0,005083	3,633723
Within Groups	0,78	16	0,04875			
Total	1,509474	18				

Fig. 4 Example B (ANOVA) in MS Excel

ANOVA

consumption

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	,729	2	,365	7,482	,005
Within Groups	,780	16	,049		
Total	1,509	18			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: consumption
Scheffe

(I) grouping	(J) grouping	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
1,00	2,00	-,20000	,11802	,267	-,5182	,1182
	3,00	-,50000*	,12928	,005	-,8485	-,1515
2,00	1,00	,20000	,11802	,267	-,1182	,5182
	3,00	-,30000	,12928	,098	-,6485	,0485
3,00	1,00	,50000*	,12928	,005	,1515	,8485
	2,00	,30000	,12928	,098	-,0485	,6485

*. The mean difference is significant at the .05 level.

Fig. 5 Example B (ANOVA) in SPSS

Variable	Analysis of Variance (Spreadsheet1) Marked effects are significant at p < ,05000						
	SS Effect	df Effect	MS Effect	SS Error	df Error	MS Error	F	p
Var1	0,729474	2	0,364737	0,780000	16	0,048750	7,481781	0,005083

Fig. 6 Example B (ANOVA) in STATISTICA

VII. EXPERIENCE WITH THE WEBSITE

In spite of this web-site being only in experimental operation, we used it for teaching Statistics at the Faculty of Education at the Charles University and Methods of Data Processing at the Czech University of Life Sciences. We would like to share some experience of its usage.

The first parts of lectures consider the basic parameter estimates, their properties. When we get to the hypothesis testing the web-site is presented to the students. We perform all the calculations by means of it and let students become familiar with it. At first, students enter various ranges of data and prove that the web-site works. They submit various simple

data and try to compare the results with manually calculated numbers and prove the presented formulas. Then they come up with their own ideas, e.g. what happens when all data cases are the same or when cases are very different, in other words when data have small or large dispersion, what happens if the range is increasing etc. They can discover some rules themselves, which is the best way of the learning. The experience with the web-site is outstanding. Students can focus on the main ideas of hypothesis testing; they can change data during their work and compare results, performing various experiments with them. Last but not least, the web-site can help the teacher to quickly create statistical examples during the preparation for teaching.

Fig. 7 Example B (ANOVA) on the website

VIII. CONCLUSION

The presented web-site is not in its final state. It can be improved by a combination of JavaScript commands or other web applications and tools. Significant improvement would be reached if we find some utility which will enable to get the values of distribution functions or the critical values of basic distributions for arbitrary arguments.

The presented web-site will serve as a significant aid in the teaching of statistics at universities.

REFERENCES

[1] D. R. Cox, E. J. Snell, *Applied Statistics: Principles and Examples*. Boca Raton: Chapnam & Hall/CRC, 1981.
[2] K. A. Carlson, J. R. Winquist, *An Introduction to Statistics: An Active Learning Approachs*. Thousand Oaks, Calif.: SAGE Publish., 2014.
[3] A. Grafen, R. Hails, *Modern Statistics for the Life Sciences*. Oxford: Oxford University Press, 2002.
[4] R. V. Hogg, J. W. McKean, A. T. Craig, *Introduction to Mathematical Statistics*. Upper Side River, New Jersey: Pearson Prentice Hall, 2005.
[5] E. Kaznowska, J. Rogers, A. Usher, "The State of E-Learning in Canadian Universities, 2011: If Students Are Digital Natives, Why Don't They Like E-Learning?," Toronto: Higher Education Strategy Associates, 2011.
[6] E. L. Lehmann, J. P. Romano, *Testing Statistical Hypotheses*. New York: Springer-Verlag, 2005.
[7] J. P. Marques de Sá, *Applied Statistics Using SPSS, STATISTICA, MATLAB and R*. Berlin: Springer-Verlag, 2007.
[8] S. Rajpal, S. Singh, A. Bhardwaj, A. Mittal, "E-Learning Revolution: Status of Educational Programs in India," in *Proc. International Multiconf. Engineer and Computer Scientists*, Hong-Kong.