MATHEMATICAL LOGIC AND STATISTICAL OR STOCHASTICAL WAYS OF THINKING: AN EDUCATIONAL POINT OF VIEW

Ödön Vancsó
Institute of Mathematics, ELTE, University of Budapest, Hungary
vancso.odon@gmail.com

At least partially, statistical thinking is based on logic different from classical logic, which causes many troubles for students. To implement didactical considerations on this issue into teaching seems to be advantageous. Some examples will be outlined to illustrate the differences in logic and how this could be transformed to actual teaching. For our school experiments we implement analogies throughout and also examples from physics teaching.

FAVOURABLE RELATION AND LOGICAL IMPLICATION

The favourable concept

In order to argue how a different logic is used in stochastical thinking contrasted to classical logic we briefly analyze a special relation between events, named "favourable relation", which was introduced by Chung (1942). The idea to use this relation in order to understand some probabilistic and statistical paradoxes goes back to Falk and Bar-Hillel (1983).

Case	Sign	Description
(1) $P(B \mid A) > P(B)$	$A \upharpoonleft B$	*A* favours *B* or *A* influences *B* positively
(2) $P(B \mid A) = P(B)$	$A \perp B$	events *A* and *B* are stochastically independent
(3) $P(B \mid A) < P(B)$	$A \downharpoonleft B$	*A* does not favour *B* or *A* influences *B* negatively

This relation can be considered as a weakened form of implication, see Vancsó (2009):

- Probabilistically taken, *A* implies *B* logically means if you presume (or imagine) that *A* has (fictionally) happened, then the probability that *B* will happen is 1 (true).
- Connected to this is the so-called favourable relation: *A* favours *B* does not mean that *B* is true if *A* (fictionally) happens; but *B* will become more probable if *A* has occurred compared to the case when *A* has not occurred.

Some features differing the favourable relation from logic implication

The logical implication follows some routine rules; for example:

- *Asymmetry*: $A \Rightarrow B \wedge B \Rightarrow A$ then $A \Leftrightarrow B$; as not all statements are equivalent, implication is *not* symmetric. For two non-equivalent events, it holds $(A \Rightarrow B) \Rightarrow \neg(B \Rightarrow A)$, i. e., the logical implication is *asymmetric*.
- *Transitivity*: $A \Rightarrow B$ and $B \Rightarrow C$ then $A \Rightarrow C$ is also true, hence the implication is transitive.

Such relations are deeply imprinted in our mind from early childhood and in primary and secondary school. It is very surprising that neither of these rules is valid for the favourable relation:

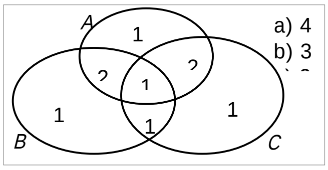
- It holds $A \upharpoonleft B$ then $B \upharpoonleft A$ the symmetry is true for all three versions of influence; i. e., the favourable relation is symmetric.
- For the transitivity, there is no general rule; sometimes it is true that $A \upharpoonleft B$ and $B \upharpoonleft C$ implies $A \upharpoonleft C$ but sometimes this does not hold.

Advantages of the favourable relation are: (i) Students become familiar with conditional probabilities and their counterintuitive features; (ii) it allows an intuitive check of calculations. A lot of paradoxes may be clarified by using special properties of this relation, which differentiate it from classical implication, see Vancsó (2009). Other rules of implication were compared to the favourable relation in Borovcnik (1992).

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

Ex. 1: If $A \Rightarrow B$ and $A \Rightarrow C$ then $A \Rightarrow (B \cap C)$.

Such a rule does not hold for the favourable relation; three different cases are represented in Figure 1, in case a) $\overline{A \cup B \cup C}$ has 4 points, in b) 3, in c) 2 points. Figure 1.



in all cases		case a)	b)	c)	always it holds
$P(B \mid A) = \frac{1}{2}$	$P(B)$	$= \frac{5}{13} < \frac{1}{2}$	$= \frac{5}{12} < \frac{1}{2}$	$= \frac{5}{11} < \frac{1}{2}$	$A \uparrow B$
$P(C \mid A) = \frac{1}{2}$	$P(C)$	$= \frac{5}{13} < \frac{1}{2}$	$= \frac{5}{12} < \frac{1}{2}$	$= \frac{5}{11} < \frac{1}{2}$	$A \uparrow C$
$P(B \cap C \mid A) = \frac{1}{6}$	$P(B \cap C)$	$= \frac{2}{13} < \frac{1}{6}$	$= \frac{2}{12} = \frac{1}{6}$	$= \frac{2}{11} > \frac{1}{6}$	
		and $A \uparrow B \cap C$	but $A \perp B \cap C$	but $A \downarrow B \cap C$	

Ex. 2: (The lack of *transitivity*) can be constructed by pupils: If A = {2, 4, 6} denotes the event of even numbers, B = {2, 3, 5} the primes, and C = {4} the numbers divisible by 4, then

$A \downarrow B$ as $\frac{1}{2} = P(B) > P(B \mid A) = \frac{1}{3}$ and $B \downarrow C$ as $\frac{1}{6} = P(C) > P(C \mid B) = 0$

but $A \uparrow C$ as $\frac{1}{6} = P(C) < P(C \mid A) = \frac{1}{3}$, which is opposite as “expected” by transitivity.

Explaining paradoxes by the favourable relation

Nemetz (1984) worked out many probability games but did not recognize the potential of the favourable relation to explain phenomena with conditional probabilities. For further examples illustrating the peculiarities of this relation, see Borovcnik (1992). Finally, we use it to enhance Simpson’s paradox.

Ex. 3: Compare mortality rates in Mexico and Sweden. In total, Sweden’s mortality rate is 5.5‰ higher than Mexico’s. In each of the age groups, however, mortality in Sweden is smaller.

	B_0	B_1	A_0	A_1	V_0	V_1	$k = V_1 - V_0$
Age	Population in millions		Number of deaths		Mortality rates ‰		Difference ‰
	Mexico	Sweden	Mexico	Sweden	Mexico	Sweden	
– 14	33.68	1.53	110,471	904	3.3	0.6	– 2.7
15 – 59	53.01	5.17	140,238	9,674	2.7	1.9	– 0.8
60 – 69	4.74	1.12	61,826	13,751	13.1	12.3	– 0.8
70 –	1.40	0.95	133,913	66,001	95.7	69.5	– 26.2
Total	93.01	8.77	446,448	90,330	4.8	10.3	+ 5.5

One explanation of this phenomenon was proposed by J. Kőrősy (1873), which is named as *standardization* in demography.

$$\overline{V_0} = \frac{\sum B_0 V_0}{\sum B_0} = \frac{33.68 \cdot 3.3 + \ldots}{93.01} = 4.8‰ \qquad \overline{V_1} = \frac{\sum B_1 V_1}{\sum B_1} = \frac{1.53 \cdot 0.6 + \ldots}{8.77} = 10.3‰$$

$$K = \overline{V_1} - \overline{V_0} = 10.3‰ - 4.8‰ = 5.5‰$$

According to this method, we have to proceed as follows: If the mortality rates of Mexico are weighted by the age distribution of the population of Sweden, we get:

$$\overline{V}_{0,\, st\ with\ B_1} = \frac{\sum B_1 V_0}{\sum B_1} = \frac{1.53 \cdot 3.3 + \ldots}{8.77} = 14.2‰$$

$$K' = \overline{V_1} - \overline{V}_{0,\, st\ with\ B_1} = \frac{\sum B_1 V_1}{\sum B_1} - \frac{\sum B_1 V_0}{\sum B_1} = \frac{\sum B_1 (V_1 - V_0)}{\sum B_1} = \frac{\sum B_1 \cdot k}{\sum B_1} = 10.3‰ - 14.2‰ = -3.9‰$$

If the population combinations were the same in the two countries and only the mortality rates were different then this rate in Sweden would be lower with 3.9‰. Reasons for the difference are:

- difference of mortality rates within age-groups (difference of ratio of parts)
- different population combination (combination of B is different)

The other type of standardization fixes mortality of “0”:

$$K'' = \overline{V}_{1,\, st\ with\ V_0} - \overline{V_0} = \frac{\sum B_1 V_0}{\sum B_1} - \frac{\sum B_0 V_0}{\sum B_0} = 14.2‰ - 4.8‰ = 9.4‰$$

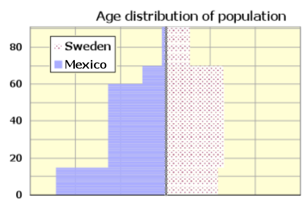


Figure 2.

If mortality rates were the same in the two countries and only the distribution of age were different, mortality in Sweden would be higher by 9.4‰. The connection between the two components is:

$$K = K' + K'' = -3.9‰ + 9.4‰ = 5.5‰$$

The situation can be described by the favourable relation as V. Bakos proposed: In the course of standardization, the entities to be compared are denoted by B and $\overline{B}$, they are grouped by the assumptions C_i ($\bigcup C_i$ denotes the whole population). The frequencies of all categories correspond to probabilities of events, for the details, see the table.

Subsets		"events"	Frequencies	Probabilities
Populations 0 and 1		B event; $\overline{B}$ complement	$\sum B_0$ $\sum B_1$	$P(B)$ $P(\overline{B})$
C_i parts of the populations (groups)		C_i as partition of Ω ($C_i \cap C_j = \varnothing$, $i \neq j$, $\cup C_i = \Omega$); C_i are "assumptions"		
C_i in 0 and 1		$B \cap C_i$ $\overline{B} \cap C_i$.	$B_0 = B_{0i}$ $B_1 = B_{1i}$	$P(B \cap C_i)$ $P(\overline{B} \cap C_i)$
Characteristic investigated (mortality)	in C_i	$A \cap B \cap C_i$ $A \cap \overline{B} \cap C_i$	$A_0 = A_{0i}$ $A_1 = A_{1i}$	$P(A \cap B \cap C_i)$ $P(A \cap \overline{B} \cap C_i)$
		$A \mid B \cap C_i$ $A \mid \overline{B} \cap C_i$	$V_0 = V_{0i} = \frac{A_0}{B_0}$; $V_1 = V_{1i} = \frac{A_1}{B_1}$	$\frac{P(A \cap B \cap C_i)}{P(B \cap C_i)}$; $\frac{P(A \cap \overline{B} \cap C_i)}{P(\overline{B} \cap C_i)}$
	in 0, 1	$A \mid B$; $A \mid \overline{B}$	$\overline{V_0} = \frac{\sum A_0}{\sum B_0}$; $\overline{V_1} = \frac{\sum A_1}{\sum B_1}$	$\frac{P(A \cap B)}{P(B)}$; $\frac{P(A \cap \overline{B})}{P(\overline{B})}$

Lemma: i) $B \uparrow A$ means that the conditional probability of A supposed B is higher compared to the unconditional, i. e., $P(A \mid B) > P(A)$. This may be characterized by $P(A \mid B) > P(A \mid \overline{B})$ ("If B favours A then $\overline{B}$ negatively influences A").

ii) Denoting $P_B(A) := P(A \mid B)$, it holds: $P(A \mid B \cap C_i) = P_{C_i}(A \mid B) = P(A \mid B \big| C_i)$, i. e., "combined conditioning events may be conceived as hierarchical conditioning".

Def.: B favours A conditional to C_i, written as $(B \uparrow A) \big| C_i$, is defined as $P_{C_i}(A \mid B) > P_{C_i}(A)$.

With these notations, the paradox may be described as follows:

- In each age group C_i, the mortality satisfies $k = V_1 - V_0 < 0$ ("mortality in country 1 = Sweden) is smaller than in country 0 (Mexico). In probabilities, this means: $P(A \mid \overline{B} \cap C_i) < P(A \mid B \cap C_i)$, or $P_{C_i}(A \mid \overline{B}) < P_{C_i}(A \mid B)$, i. e. (see Lemma), $(B \uparrow A) \big| C_i$.
- In the whole population, $K = \overline{V_1} - \overline{V_0} > 0$, mortality in 1 is higher than in 0. Read as probabilities, this means $P(A \mid \overline{B}) > P(A \mid B)$, i. e., $B \downarrow A$.

Can it be that for all cases C_i, which form a partition of the universe Ω, that it holds $(B \uparrow A) \mid C_1$, ... , $(B \uparrow A) \mid C_r$ and yet $B \downarrow A$ for the universe? This seems to be impossible. The paradox comes from a domination of our thinking by logic where the following is true:

If $C_1, C_2, ..., C_r$ form a partition of the universe Ω then it holds

$\{ (A \Rightarrow B) \mid C_1,\ (A \Rightarrow B) \mid C_2, \ldots\ (A \Rightarrow B) \mid C_r$, then $A \Rightarrow B \}$

This "rule" is used by the principle of proof by case discrimination. If one can prove $A \Rightarrow B$ in each of the cases C_i and the cases cover all possibilities, then the proof is complete. Such a rule, however, can not be transferred to probabilities and the favourable relation as we have seen by Simpson's paradox above. As Borovcnik (1992, pp. 211) formulated: in the case of the favourable relation a boundary condition can not be lifted into assumptions opposed to implication.

CONNECTING TEACHING OF STATISTICS AND PROBABILITY TO MODERN PHYSICS

Incentives from modern physics teaching

There was a didactical stream in physics (in parallel to the New Math) to embed modern ideas of science in physics at school. Kuhn (1962), in his paradigm theory, elaborated on the psychological and cognitive obstacles that humans encounter when they should undergo revolutionary changes. That makes it even more important to confront our young students early enough with the new concepts in physics otherwise they might "refute" to change their concepts already acquired later. This thesis lead to a didactical stream in Hungary; Prof. G. Marx and his group tried to change teaching physics in the late primary and secondary schools accordingly (see Tóth and Marx 1980-82). Ideas from Károlyházi (1981) remind us to similarities between teaching physics and statistics: "The reason lies in the special nature of the difficulty with quantum mechanics. The essence of this difficulty is not the complexity of things; it is rather that *we hear about something simply senseless* (italics by the author)." Later he writes more explicitly "the trouble is not that the electron is different from whatever we have seen earlier but *this existence appears a logical impossibility*." We remember that the cognitive situation is quite comparable to e.g., Simpson's paradox. The situation involved seems to be *impossible*. The question is why? Reading Károlyházi further: "Our tactics, on the contrary, is to emphasize the outlined nature of difficulty." This is also *our* starting point. "The very first question is: How can empirical facts appear not only awkward but logically impossible? Answer: Some time, way back in our past, perhaps while being sucklings, we have created a false picture about a phenomenon so basic that we are unable to remember now the process of getting acquainted with this phenomenon and we feel the false picture so natural that we subconsciously smuggle it into our thought-line. This lays the foundation for the contradiction."

We will not pursue this article further because it deals with quantum mechanics and our topic is statistics. But the essence of these thoughts can help us to plan a strategy in teaching statistics using similarities between these two fields. In our case, the earlier experience is the classical implication, which has special rules imprinted in students' minds by continuous learning. In a new case where these rules fail, it causes troubles to us. We have to confront our students with probabilistic thinking as early as possible before the classical logic has fixed their mind.

Another physicist, Weisskopf (1981) focused on the importance of conceptual learning: "The whole conceptual framework of quantum mechanics upon which modern physical understanding of the properties of matter is based. The difficulties encountered here are not necessarily mathematical: they are conceptual." We agree that conceptual teaching and learning is very important. In the 1970's, G. Marx a physics professor of Eötvös University tried to focus on it. His main idea was to teach physics by modelling; students should get experience with the models used. The corn concept is to confront students with modern physics earlier than usual. See Marx and Tóth (1991) for an overview about this topic.

To sum up: reading an interesting teaching idea in physics we may get encouraged to introduce statistics and probability in schools as early as possible. We have to prepare careful teaching materials from the very beginning in order to offer encounters with "strange or unusual appearances". One of the crucial concepts is conditional probability because it lies behind many problems and paradoxes. It should be shown that many situations in everyday life may be handled by them. Simpson's paradox seems to be contradictory in classical logic but "normal" in statistics.

A teaching experiment involving Bell inequalities

Another topic where classical logic on the one hand and probabilistic thinking on the other may get in conflict is Bell inequalities, which embody that there is no classical cause and effect law in quantum physics. We can imagine only two types of causal connections between two events *A* and *B:* Either one of them follows from the other, or both have a third event as common root, which causes both of them. There is a quantum phenomenon which shows that there should be another, *third type - causal connection*, between *A* and *B*. The situation was modelled by an experiment (see Hraskó, 1984), which focused on Bell inequalities; the mathematical level is so elementary that it can be handled in schools as the author tried out in a secondary school.

Ex. 4: Three persons are taking part in an imaginary experiment. All of them are in different rooms, *C* in the middle room, *A* in the left, *B* in the right room. See Figure 3. On the desk of *C* there are a plenty of identical closed big envelopes. All contain two identical middle size envelopes. At the end, all middle size envelopes contain three little envelopes, which are numbered from 1 to 3 (see Figure 4). Opening an arbitrary little envelope, this contains a *red* or a *blue* disc. The colour distribution of these discs is the subject of our analysis. There are rules for opening envelopes. In the middle room, *C* chooses a big envelope, opens it and of the two middle size envelopes he sends one to *A* and the other to *B*. Both of them open the envelope and choose one little envelope and open it; the other two little envelopes are destroyed. *A* and *B* note the number of the chosen little envelope and the colour of the disc. Then *C* chooses another big envelope and the process will be repeated. After examining a lot of envelopes, the partners finish the experiment and compare the lists of *A* and *B*.

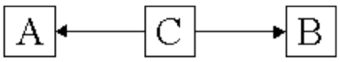


Figure 3.

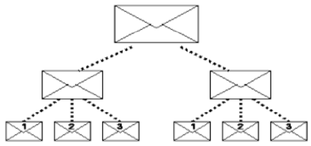


Figure 4.

List of *A*	List of *B*
1B	2R
3R	1B
2B	2R
2R	3R
1R	1B
2B	1B

The first rows can be read in the table next to Figure 4. If the number is the same in one row, the colour is always different (occurs in third and fifth row). The distribution of the colour in the little envelopes is not totally due to chance: In the same big envelope the little envelopes with the same number always contain discs with different colour. Consequently, if we see the two lists, we see a connection between them. Which type of correlation could be between the two lists? It seems to be common reason type correlation because information was not flowing directly between *A* and *B*. Easier to say the unknown creator of the big envelopes took discs with different colour in the little envelopes with the same number very precisely.

But *is it possible to read from the data* of the table, which contains only the colour of discs after opening the envelopes, that they *have the colour determined before opening*? It is very surprising but *the answer is yes*. In the table, 36 different row types occur (2 colours and 3 numbers give 2·3=6, for two columns 6^2). From these, there are 12 cases where the numbers are the same but from these only 6 occur; this reduces to 15 cases. Let denote by *N*(*i*R, *j*B) the number of the rows which contain *i*R in one and *j*B in the other column. If the table contains many rows, the probability can be estimated from the empirical relative frequencies. In such a way, we estimate

$$p(1R,2R) = \frac{N(1R,2R)}{N} \quad (N \text{ denotes the number of all rows}).$$

15 cases
(1R, 1B)
(1R, 2B)
(1R, 2R)
(1R, 3B)
(1R, 3R)
(1B, 2B)
(1B, 3B)
(2B, 3B)
(2R, 1B)
(2R, 3B)
(2R, 2B)
(2R, 3R)
(3R, 1B)
(3R, 2B)
(3R, 3B)

The 15 probabilities are estimated from the results; to these the 6 known probabilities are added: $p(iR,iR) = p(iB,iB) = 0, \quad i = 1, 2, 3$

Type of the big envelope	Content of the one middle size envelope	Content of the other middle sized envelope
a	1R, 2R, 3R	1B, 2B, 3B
b	1R, 2R, 3B	1B, 2B, 3R
c	1R, 2B, 3B	1B, 2R, 3R
d	1R, 2B, 3R	1B, 2R, 3B

These equalities express the very closed correlation between the two events column. If we add the *assumption* that all discs have a colour from the beginning, then there must be some inequalities among the 15+6 = 21 empirical probabilities, which are named Bell inequalities after their discoverer. Since the information is kept carrying by the colour of discs, we have to convince ourselves that the discs have a colour in the envelope *before* it is opened. If the colour of the disc were determined by e. g., the meeting with air then the information of the colour would not be carried by the envelope. Examine the situation whether we could be convinced about the determined colour before the letter is opened. Because of the rules of the game, there are only four types of big envelopes with probabilities $w_{i,}$. Bell inequalities are based on the positivity of w_i using the probabilities *p*. One typical is: (*) $p(1R,\, 2R) + p(2R,\, 3R) \geq p(1R,\, 3R)$

Such an inequality is proven by expressing the *p*-probabilities by w_i. *p*(1R, 2R) can be expressed by w_i: 1R, 2R is only types *c* or *d*, the probability is $w_c + w_d$ and the probability that *A* opens the first and *B* opens the second little envelope is $(\frac{1}{3})^2 = \frac{1}{9}$. From that we have:

$$(1)\quad p(1R,2R) = \tfrac{1}{9}(w_c + w_d)$$

All the 15 different probabilities of *p*(*i*R, *j*B) (with the restrictions, see the table of cases) can be written by four different w_i probabilities. From these, there are only three independent using $w_a + w_b + w_c + w_d = 1$. Hence, many constraints have to be fulfilled among the *p*-probabilities. The other two p-probabilities may be expressed by w_i in the following way:

(2) $p(2R, 3R) = \frac{1}{9}(w_b + w_d)$ (3) $p(1R, 3R) = \frac{1}{9}(w_b + w_c)$

If we substitute (1), (2) and (3) in (*), we get the proof for (*). This inequality (*) unifies two components: the *empirical fact* about a colour of discs in the little envelopes and the *assumption* that the colour of the discs was determined as it was put in the envelope. If the probabilities p (as estimated from the empirical lists) do not satisfy a Bell inequality, then one of these pieces of information is not valid. The first is “proven” by the experiment; only the second – our assumption – could be false. In this case the correlation is not based on common reason. But we are able to eliminate direct information from *A* to *B* (“we take *A* very far from *B*“). It means if the inequality is violated then a special new type of correlation should exist.

Conflicts with Bell inequalities as starter for quantum logic

In quantum physics it occurs that these rules are unavoidable, we can not measure complementary quantities. If we measure the spin of a proton and neutron, which form together a deuteron atom, the sum of spins amounts to 0. After the fissure of this deuteron two components are flying in opposite directions with the same velocity. In two laboratories the spin component of the proton and the neutron are measured in only one of three spatial directions. The results have to be different. Making many experiments we see that the Bell inequalities are not valid. It means that in quantum mechanics there is another type of correlation as we have learned it. This model of a real phenomenon, which can be taught in school, helps us to understand its paradox characters (see Shimony 2006 for a complex summary about this topic).

Conclusions

We have seen some important differences between logical implication and the favourable relation. Many “contradictions” would be avoided if the classical thinking has not yet imprinted our mind; therefore conditional probability has to be taught as early as possible. Nobody fails to notice the lack of transitivity in a football league but in mathematics it is really problematic because in this field we have been involved so much with transitive relations (<, =, similarity, etc.) that by the end we can not imagine such a situation any more where transitivity fails.

From many points of view, the situation in teaching statistics or modern physics is similar. Quantum logic *is* probability logic. Probability and statistics play a central role in physics and to extend such links would promote a cooperation and collaboration useful for both sides.

REFERENCES

Borovcnik, M. (1992). *Stochastik im Wechselspiel von Intuitionen und Mathematik*. Mannheim: BI.

Chung, K. L. (1942). On mutually favourable events, *Annals of Mathematical Statistics 13*, 338-349.

Falk, R., & Bar-Hillel, M. (1983). Probabilistic Dependence between Events, *Two-Year College-Mathematics Journal, 14,* 240-247.

Hraskó, P. (1984). A Bell-egyenlőtlenség. *Fizikai Szemle 1984* (7), 257-264. (in Hungarian) Retrieved Dec., 30, 2009, from www.hrasko.com/xaknak.php?docid=10015.

Károlyházi, F. (1981). How to get a right feeling about quantum mechanics? In *Quantum mechanics in the school*. Budapest: Eötvös Physical Society.

Kőrősy, J. (1873). *Plan einer Mortalitätsstatistik für Grosstädte*. Wien: Universitätsverlag.

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Marx, G. and Tóth, E. (1991). Models in Science education. *Educating for an unknown future* Veszprém: National Centre for Educational Technology.

Nemetz, T. (1984). *Probability for special mathematics classes.* Budapest: Nat. Schoolbook Publ.

Shimony, A. (2006). Bell’s theorem. *Stanford Academy.* www.stanford.com (retrieved Dec., 30, 2009).

Tóth, E., & Marx, G. (1980-82). *Physics for grades 9-12*. Budapest: National Schoolbook Publ.

Vancsó, Ö. (2009). Parallel discussion of classical and Bayesian ways as an introduction to statistical inference. *IEJME, 4*(3), 269-309 Online: www.iejme.com.

Weisskopf, V. (1981). Is Physics human? In *Quantum mechanics in the school*. Budapest: Eötvös Physical Society.