Chapter 40
Popularization of Probability Theory and Statistics in School Through Intellectual Competitions

Ivan R. Vysotskiy

Abstract Since 2004, in accordance with the Federal Educational Standards, probability theory and statistics has been included into teaching practice in Russian schools. This paper focuses on one form of this work: organization of intellectual competitions on probability theory and statistics for school students. Since 2008, the Moscow Center for Continuous Mathematical Education has conducted the Internet Olympiad for students in school years 6–11. In addition to the traditional problems, participants are offered a choice to write an essay on a proposed topic. This article attempts to classify those topics and highlight the most popular ones among the students. In addition, this paper makes a short overview of selected problems that from the organizers’ point of view represent promising and prospective trends in the teaching of probability and statistics at school. The article is addressed to education specialists, teachers, and researchers who specialize in probability theory and statistics.

Keywords Probability · School math education · Olympiad on probability · Math intellectual competition · Moscow center for continuous mathematical education

40.1 Introduction

In 2004, elements of probability theory and statistics were introduced into the school mathematics curricula in Russia in accordance with the federal educational standards. Since 2012, problems on probability and data representation have been included in the Unified State Exam in mathematics. The Concept of Development of Mathematical Education in Russia was approved by the government in December 2013. It states that probability theory and statistics are important sections of school mathematics. In 2015, the Federal Exemplary Curricula were developed,
where probability and statistics appear as complete sections that determine content for each educational level.

Meanwhile, educational practices have caused serious difficulties that have been inescapable when forming a new school subject that is completely different from the traditional courses that are being taught in universities. Difficulties in the preparation of teachers have followed from these difficulties.

This paper focuses on only one of the popularization dimensions aimed at the formation of public inquiry into the field of mathematics: the methodology and practice of intellectual competitions for school children on probability theory and statistics. The example considered was the Olympiad that has been held by the Moscow Center for Continuous Mathematical Education (MCCME) since 2008.

The Olympiad Organizing Committee is ready to cooperate with colleagues from all countries. In particular, we can provide English versions of all Olympiad materials.

### 40.2 Background

We often hear from mathematicians that because probability theory and statistics are too difficult, they should not be taught in school. This opinion has grown out of complications that follow studying probability theory in universities, which traditionally is deductive and based on combinatorics and wide knowledge of calculus. A combinatorial approach to probability theory is typical for many generations and has grown from the Soviet period, when probability theory was torn out of statistics upon being announced as “a social science.” In fact, combinatorics is not directly related to basic ideas of statistics and probability.\(^1\) This is just a way to enumerate elements of vast probabilistic spaces and prove theorems. Experience and intuition are primary, and no combinatorial tricks are useful without them. One should meaningfully consider chances of events, especially unlikely events that play a significant role in daily life.

The problem is that events are less obvious than figures or numbers, while concepts of chance and volatility are not as intuitive as length, area, or volume. An event and its chances make special types of abstract objects and their formalization into mathematical notions is much more complex than a formalization of a picture (transition to geometry) or a quantity (to arithmetic or algebra).

The second problem is that for the majority of children, the concept of volatility and instability of events is alien to them until a certain age. At what age a child is ready to perceive changeable models and determine which models they should be has yet to be discovered. In the Soviet period, the science of the laws of cognitive

\(^{1}\)In talking about the probability theory in school, we often omit “theory” to simplify the text and follow the tradition that has been formed in Russian educational terminology.
activity was destroyed (Petrovsky 1991). Its place was taken by the paradigm of “a clear sheet of paper.”

However, if in early childhood the rejection of variability possibly serves as a defensive mechanism that simplifies adaptation to social and natural conditions and minimizes the number of necessary rules of behavior, then later an absence of general ideas of randomness and volatility starts to hinder efforts to understand abstract concepts of probability and statistics. Study of statistics in adulthood does not improve this situation (Kakihana and Watanabe 2013).

The first thing to take care of is to make basic concepts of probability and statistics clear and familiar to math teachers who have difficulty in the transition from teaching abstract mathematical facts to applying mathematical concepts and laws to the solution of practical problems.

In addition, one of the most important aspects of education is the popularization of knowledge. If algebra, geometry, and other sections of traditional school mathematics show no shortage of additional popular scientific and popular literature for adults and children of all ages, an analysis of the situation in the area of probability theory has shown a distinct lack of a popular literature and other forms of popularization (Bunimovich et al. 2009a, b). A number of popular books for school children were published in the Soviet Union, mainly in the 1950s through the 1980s (see, for example, Kolmogorov et al. 1982; Mosteller 1985). The number of new materials appearing in Russia is vanishingly small, even if we take into account translations of foreign publications on probability and statistics for students (e.g., Chjun and Ait-Sahli 2007). At the same time, the amount of popularization literature on statistics and probability theory in the world has been increasing. This is partly due to the increasing importance of teaching probability and statistics at school and partly due to the growth of the role of stochastic methods in different sectors of the global economy (Bunimovich et al. 2009a). Popularization activities should be carried out in various forms. In addition to special literature and mathematics circles, Olympiads and other intellectual competitions of different levels for children and adults are of great value. Recent years have shown a spontaneous increase in number and quality of Internet sites dedicated to popular mathematics, in particular in the field of probability theory and statistics.

40.3 The Olympiad, Its Main Principles and Description

In 2008, the first Olympiad was administered by MCCME on the initiative of Yuri Tyurin, Alexei Makarov, and Ivan Yaschenko. In following years, many mathematicians and educators (E. Bunimovich, V. Bulychev, P. Semenov, et al.) participated in the selection and preparation of Olympiad problems. The rules of the Olympiad and its requirements for participation are simple. The Olympiad is open to everyone and is held during a calendar month. Olympiad problems are designed for students in Grades 6–11.
The Olympiad has two rounds. The first round, which is invitational, happens in schools. In 2015, the number of participants was 2450 and in 2016 it was 2890. In 2016, an intramural final round was added that consisted of two contests: first, a competition involving a statistical experiment and, second, problem solving.

The problems in the Olympiad are designed to be recommended to students starting from a certain grade level. For example, a task that can be solved using only intuitive ideas, finite enumerations, and classical probability definitions is recommended for students in Grade 6 and older. If a solution requires the use of simple transformations within the algebra of events, the task is recommended for students in Grade 7 and older. A problem involving the characteristics of random variables will be offered to students in Grade 8 and older. There is no upper limit to the ‘age’ of a problem. Age differentiation occurs at the stage of grading solutions and awarding the winners.

Olympiad problems are placed on the website for free access for about a month. According to the rules of the Olympiad, participants can use any help, reference books, etc. Only “openly hiring adult labor” is considered non-sporting and is discouraged.

Materials from previous years are published on the Olympiad website in the Archive section at http://ptlab.mccme.ru/olympiad.

Olympiads from 2008 to 2017 have been published. Articles dedicated to this competition have been published in (Vysotskiy 2012; Vysotskiy et al. 2009).

### 40.4 Essay Tasks

A distinctive feature of the Olympiad is that it also includes essay tasks in addition to problems. The essays are evaluated separately, regardless of the age. The essays turned out to be important and most attractive part of the Olympiad.

Participants are required to analyze the offered situation and write a short essay on a given topic. These essays immerse students in uncertain situations that require imagination and activities involving estimation that take into account real limitations and the nature of the data. Actions in an uncertain situation play a crucial role in the formation of the statistical and mathematical culture of students, because instead of performing the steps of a known or studied algorithm, Olympiad participants become researchers who plan experiments themselves. Students have to determine important and unimportant factors of random experiments, interpret the results, invent a method to describe the data, and form a hypothesis. None of the proposed essays requires students to check the formulated hypothesis, as the mathematical tools available to school students are clearly insufficient for that.

In some situations, the creation of a hypothesis itself is a very complicated task that requires a high culture of regulatory activity from a student. Moreover, situations often arise where the number of possible plausible hypotheses is large. In such cases, the authors try to repose the situation in such way that either it contains
a hypothesis formulated explicitly or the task implicitly directs the students’ actions (see two examples below). The most important part of the essay task for a student is to search for data with the aim of finding a pattern or checking the data’s correspondence with certain assumptions. Some tasks require students to collect data independently through short surveys. The others demand an independent search on the Internet.

Some essay tasks are designed to generate critical scientific thinking in students, immersing students in situations where the arguments provided deliberately contain an inaccuracy of some kind, an error, or an unreasonable conclusion. Students are invited to understand the shortcomings of the study and make research steps. In my experience, tasks of this type appeared to be the most attractive to the participants. Among the submitted essays there have been very bright and original works, which we placed on the Olympiad website in the Solutions section.

Below we provide several essay tasks from previous years.

All essays can be classified by their educational objectives and the methods used to perform them:

1. **Checking a statement** using either independently collected data or raw data provided in the essay task. In this case, the essay usually offers to an Olympiad participant a partially proposed algorithm of actions.

2. **A study that may require a survey** of the student’s classmates, friends, and parents followed by data processing with the formulation of a plausible hypothesis or the refutation of an implausible one. A student’s ability to organize collected data, present it in the most appropriate way, and produce hypotheses are important here.

3. **A search for a statistical method** for solving a problem (see Essays 1 and 2). This is aimed at developing constructive thinking and skills. An unbounded search appears very attractive to students. An essay about the estimation of a number of people in a crowd was the most frequently selected among participants of all age groups in 2013. None used a method that the authors of the problem considered the most natural; instead we received many original ideas.

4. **A search of an error** in a complex and extensive reasoning presented in the formulation of the task (see, e.g., Essay 6). Solving tasks of this type helps to develop a critically destructive way of thinking, which is an integral part of intellectual culture. Speaking specifically on the implementation of an essay about the relationship between air humidity and levels of snowfall, we note that many students succeeded in noticing flaws in the author’s arguments, but the main problem—the unsuitability of the linear regression for the case—was not mentioned by any of them.

This shows that the basic ideas of statistics stay on the border between intuitive and conscious.

**Examples of essay tasks**

1. **The number of people in a crowd (2012).** The photo on Fig. 40.1 shows a crowd of people. How one can estimate (approximately calculate) the number of
people in this crowd? Try to develop an appropriate method, use it and make such estimation. Describe your method in detail, explain why it works correctly and how to use it, and what you got using it. We are looking forward to your answer: How many people are there in the photo?

2. **Which way is faster (2014)?** Is it true that an aircraft takes a different amount of time when it flies east and west? Is it always so? Go to the site of a large airport or a major airline; it is even better if you consider several airlines or airports. Select flights from east to west and vice versa. Collect and process the necessary information. How different are the durations of flights one way and the other? Does it depend on the distance? You need to come up with a statistical measure that describes the difference. Is it stable? The difficulty is that a mere difference between the time there and the time back does not give us much: We have to take into account not only very long flights but also relatively short ones. If such a difference does exist, what causes it? Is it possible to estimate consistency and power of this amazing factor? Are east and west really guilty? Maybe a similar pattern can be observed with other flights, for example, from north to south and back? Many questions arise. Try to locate, describe, and analyze data and use your imagination.

3. **Haga’s problem (2011).** Professor Kazuo Haga from the University of Tsukuba is the inventor of origamics (geometrical origami). Once he posed an interesting question: A paper square is divided into light and dark parts by four semicircles (Fig. 40.2, left), giving a graceful ornament that resembles a flower.

   If we choose a point in a light part and then make folds so that all vertices meet this point, we get a pentagon (Fig. 40.2, right). So the union of light parts is a “domain
of pentagons” or simply a “5-domain.” The dark parts make a “6-domain.” The four vertices and the center of the square give quadrangles. Professor Haga writes: “I noticed that when asked to choose a random point, the majority of people mark a point leading to a hexagon. Pentagons are much rarer. Very few choose points that make quadrangles. The question is: Is the number of those who choose a point in a certain domain proportional to the domain’s area?”

If sides of the square measure 1 unit, then each semicircle on the left figure has radius of 0.5 and its area is $\pi/8$. Therefore, the area of the petals (6-domain) equals $\pi/2 - 1 \approx 0.57$, while all the rest (5-domain) has the area of about 0.43. The difference is not too big. The ratio of those who choose points in the corresponding domains does not fit the ratio of the areas. What is the reason for this: Why are points outside the petals less attractive than points inside?

Conduct an experiment. You will need some a number of paper squares. Ask as many people as possible to mark a point on the clear square. Putting all points together on a new square gives the distribution of the points. Maybe some properties of the distribution will help to explain why “hexagon admirers” appear more often than Professor Haga could anticipate, having compared the domain areas.

4. **Height correlation (2009).** A teacher once decided to show her students that heights of boys and girls are independent random values. In order to do this, she did some research. In every class, she chose 10 girls and 10 boys at random, then composed random boy/girl pairs and wrote down their heights as $x_k$ and $y_k$ for every pair. When she was done and had put all the results on the scatter chart (Fig. 40.3), she found to her horror that all the points were grouped around a slanted line. This meant that there was an obvious correlation between heights of boys and heights of girls! How could this be?

Write a short essay in which you try to explain whether the teacher made a mistake in her findings and, if so, what her misjudgment likely was.

5. **Insurance (2010).** The insurable value of a car depends on its age. Agents of insurance company ABC estimate the depreciation in a very simple way:
Cars older than two years old lose 10% of their value yearly. Using http://www.auto.ru and other available sources, conduct research on the topic of whether the price policy of the company ABC corresponds to actual practice in the used car market. In your research, take into account that among the cars there are some vehicles whose values do not meet the average for the cars of the same model and age.

### 40.5 Problems That Require a Solution

In addition to three essays, the Olympiad traditionally includes 16 problems in statistics and probability theory. Some problems are easy and admit a very simple solution by brute force or short reasoning. More complex problems require, in addition to looking for the key to the problem, students to have the ability to perform operations on events and some knowledge of probability properties. Finally, there are some very complex problems that surrender only to those who devote enough time to thought and attempts and who have studied the literature and problems from previous years. Authors deliberately include in the Olympiad special problems whose solution requires complex transformations. Such problems are few, but they must be in an online competition, as participation in this type of problem implies that the participant has done scientific research. A few examples are listed below. We wanted to give a sample of problems that show a wide range of methods and topics used in probability. The difficulty of each problem is indicated by its number of points and recommended grade.
Examples of problems

1. **Defective coins (sixth grade and older, 1 point).** For the anniversary of the Saint Petersburg mathematical Olympiad, the mint produced three commemorative coins. One coin was made correctly, the second coin had tails on both sides, and the third coin had heads on both sides. The Director of the Mint chose one of these coins without looking, tossed it, and it came up tails. Find the probability that the second side of this coin also is a tail.

2. **Three targets.** A shooter fires on three targets as many times as needed to hit all three. The probability to hit for one shot is \( p \).
   
   (a) **(Seventh grade and older, 2 points).** Find the probability that the shooter will fire exactly five times.
   
   (b) **(Eighth grade and older, 2 points).** Find the expected value for the number of shots.

3. **Intersecting diagonals (ninth grade and older, 3 points).** In a convex polygon with an odd number of vertices equal to \( 2n + 1 \), two random diagonals are chosen independently. Find the probability that these diagonals intersect inside the polygon.

4. **Draws (ninth grade and older, 6 points).** Two hockey teams of equal strength have agreed that they will play until the total score reaches 10. Find the expected value for the number of times a draw will happen.

5. **Stunning news.** A conference is attended by 18 scientists, of whom 10 know some stunning news. During a coffee break, all scientists are randomly divided into pairs and in each pair a scientist who knows the news tells it to the other if the other did not know it yet.
   
   (a) **(Ninth grade and older, 1 points).** Find the probability that after the coffee break the number of scientists who know the news will be equal to 13.
   
   (b) **(10th grade and older, 4 points).** Find the probability that after the coffee break the numbers of scientists who know the news will be equal to 14.
   
   (c) **(Ninth grade and older, 3 points).** Denote by \( X \) the number of scientists who will know stunning news after the coffee break. Find the expected value of \( X \).

6. **Mini-Tetris.** A tall rectangle of width 2 is open at the top, and randomly oriented \( L \)-shaped triminos fall into it. (A trimino is a piece that looks like a domino piece but consists of 3 squares. There are two types of trimino, straight and \( L \)-shaped.)
   
   (a) **(Ninth grade and older, 3 points).** Let \( k \) triminos fall into the rectangle. Find the expected value of the height of the resulting polygon.
   
   (b) **(10th grade and older, 6 points).** Let 7 triminos fall. Find the probability that the resulting figure will have a height of 12.
40.6 The Statistical Experiment Contest

In 2016, the Olympiad had three rounds for the first time. In the final (intramural) round, participants were offered a topic where they were to develop a statistical experiment. Organizers took a well-known scheme of taste tests as a base for it. All participants asked to design an experiment aimed at revealing the threshold of sensitivity to sweetness using a weak aqueous solution of sugar. All participants were asked to remember the following.

1. Hygiene. It is unacceptable to have more than one person drink water from the same cup.
2. Various effects that distort the result are possible. For example, a person may be less sensitive to a less saturated sugar solution after a stronger one. How can this effect be reduced?
3. There may be many tiny factors that affect taste. Should we regard them all as non-significant? Should some of them be taken into account?
4. How to process the collected data. We do not assume a deeply scientific approach, but we hope that a proposed procedure will be convincing.

Participants presented several plans for such an experiment. After a discussion, the best-justified plan was implemented. (On Fig. 40.4, left: the course of the experiment; right: the best plan author Amelia, fifth grade, Republic of Bashkortostan.)

During the discussion about this form of work, the organization committee came to the following conclusions:

1. This experiment may become the most dynamic and exciting part of the Olympiad.
2. The topic of the experiment should be chosen carefully and designed so that learners should be able to replicate the experiment in the existing conditions within the announced time.
3. The form of the experiment should be chosen in such a way that all participants can take part together with their parents, accompanying persons, etc., regardless of age.
4. The experiment should be planned so that all participants will be busy at all times or will at least have a chance to busy themselves by performing tests, collecting statistics, processing collected data, etc. Optimally the experiment should be organized so that work is done in groups with different duties for all members in a group.

---

2Not a mistake. We have announced the Olympiad for 6–11 grades but if a fifth grader wants to take part—why not? She is a really smart girl.
40.7 Conclusion

The experience of the Olympiad on probability theory shows that despite the fact that this branch of mathematics has traditionally been considered difficult and unusual for school learning, interest in it is gradually increasing.

In 2015, the Olympiad school tour was held for the first time and was attended by 2465 students from 20 regions of Russia, and in the 2016 tour, 2859 students from 41 regions participated, as the Olympiad was joined by leading schools in many regions where such a competition had not previously been known.

Growth has also been indicated by the increasing number of requests from teachers and students to provide methodological support for teaching and learning. During the 2014–2015 study year, the methodological site for teaching of probability and statistics (http://ptlab.mccme.ru) received 876 queries. During the 2015–2016 school year (up to May 13) the number of such inquiries was 1764. The proportion of queries related to the Olympiad only (applications, rules, results, appeals, etc.) was 20% (downloads). These statistics show that the Olympiad plays an important role in the popularization of probability theory and statistics as a school subject.

Of course, the Olympiad should not be the only means of promoting this subject. In addition, there must be math circles (out-of-class activities), numerous publications on probability and statistics in teachers’ and popular magazines (such as Mathematics, Mathematics in School, Kvant, and Kvantik). At the same time, the number and variety of probabilistic problems on the national exam for primary and high school courses has increased.

Unlike the problems from the regular school course, the Olympiad tasks are much more diverse in subject matter and level of difficulty. Taking advantage of this, the Olympiad developers are gradually expanding the range of tasks and inventing new forms, some of which later will be included in school courses and methods.
References

Bunimovich, E. A., Bulychev, V. A., Vysotskiy, I. R., et al. (2009a). About probability theory and statistics in the school course. Mathematics in School, 7, 3–13.

Bunimovich, E. A., Vysotskiy, I. R., et al. (2009b). Terminology, notations and agreements in a school course of probability theory and statistics. Mathematics, 17, 13–27.

Chjun, K. L., & Ait-Sahlia, F. (2007). An elementary course of probability theory. Stochastic processes and financial mathematics. Moscow: BINOM, Laboratoriya Znanii.

Kakihana, K., & Watanabe, S. (2013). Statistics education for lifelong learning. In Proceedings of 6th East Asia Regional Conference on Mathematics Education (EARCOME6) (Vol. 3, pp. 318–322), March 17–22, 2013, Phuket, Thailand.

Kolmogorov, A. N., Zhurbenko, I. G., & Prokhorov, A. V. (1982). Introduction to probability theory. Moscow: Nauka.

Mosteller F. (1985). Fifty challenging problems in probability with solutions (3rd ed.). Moscow: Nauka, FIZMATLIT.

Petrovsky, A. V. (1991). Ban on a comprehensive study of childhood. In Repressed science (pp. 126–135). Leningrad: Nauka.

Vysotskiy, I. R. (2012). Olympiad problems on PT. In Mathematics (January issue, pp. 61–62). Pervoe Sentyabrya Publishers.

Vysotskiy, I. R., Borodkina V. V., et al. (2009). Online olympiad on probability theory and statistics. In Mathematics (Vol. 15, pp. 17–25). Pervoe Sentyabrya Publishers.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.