

Conversations with Gábor J. Székely

Yulia R. Gel, Edsel A. Peña and Huixia Judy Wang

Abstract. Gábor J. Székely was born in Budapest, Hungary on February 4, 1947. He graduated from Eötvös Loránd University (ELTE) with an M.S. degree in 1970, and a Ph.D. degree in 1971. He received his Candidate Degree from the Hungarian Academy of Sciences in 1976, and the Doctor of Science Degree (D. Sc.) from the Hungarian Academy of Sciences in 1986. Székely joined the Department of Probability Theory of ELTE in 1970. In 1989, he became the founding chair of the Department of Stochastics of the Budapest Institute of Technology (Technical University of Budapest). In 1995, he moved to the United States as a tenured full professor at Bowling Green State University (BGSU) in Bowling Green, Ohio. Before that, in 1990–1991, he was the first Lukacs Distinguished Professor at BGSU. Székely had several visiting positions, for example, at the University of Amsterdam in 1977 and at Yale University in 1989. Between 2006 and 2022, he served as a Program Director in the Statistics Program of the Division of Mathematical Sciences at the U.S. National Science Foundation. Székely has about 250 publications, including 6 books in several languages. In 1988, he received the Rollo Davidson Prize from Cambridge University, jointly with Imre Z. Ruzsa for their work on algebraic probability theory. In 2010, Székely became an Elected Fellow of the Institute of Mathematical Statistics mostly for his works dealing with physics concepts in statistics like energy statistics and distance correlation. He had the fortune to know and work with world-class mathematicians and statisticians like (in chronological order of their first meetings): P. Erdős, A. Rényi, Y. Linnik, B. de Finetti, A. N. Kolmogorov, H. Robbins, G. Pólya, L. Shepp, G. Wahba, C. R. Rao, B. Efron, P. Bickel and E. Seneta.

Key words and phrases: Distance correlation, energy of data, Lukacs Professorship, negative probability, uncertainty principle of games.

1. INTRODUCTION

Gábor J. Székely served as a Permanent Program Director of the Statistics Program in the Division of Mathematical Sciences of the U.S. National Science Foundation (NSF) from 2006 until his retirement from NSF in February 2022. The authors wanted to chronicle his colorful and interesting life, from his native Hungary to Bowling Green State University in Ohio, and then to the NSF in Virginia. What follows is based on a series of conversations, via Zoom, e-mail and in-person, between him and

the three authors, who were all then Rotating Program Directors in the Statistics Program at NSF. The virtual conversations were held intermittently during the depths of the COVID-19 pandemic when NSF personnel were working virtually, and completed in Alexandria, Virginia, in-person, near the NSF Headquarters on May 11 and 23, 2022. In the conversations that follow, **GS** refers to Gábor, **YG** refers to Yulia R. Gel, **EP** refers to Edsel Peña, **JW** refers to Judy Wang and **GPW** refers to the three authors.

2. FAMILY AND BACKGROUND

YG: Gábor, please tell us about your family, childhood and background, if possible.

GS: I know a lot about my mother's family. For many generations they lived in Mattersdorf, Burgenland, close to the present day Austrian–Hungarian border on the Austrian side, until about the end of the 19th century when they moved to Hungary and eventually settled in Budapest around 1930. They then never left Budapest. The

Yulia R. Gel is Professor, Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75080, USA (e-mail: ygl@utdallas.edu). Edsel A. Peña is Professor, Department of Statistics, University of South Carolina, Columbia, South Carolina 29208, USA (e-mail: pena@stat.sc.edu). Huixia Judy Wang is Professor and Chair, Department of Statistics, George Washington University, Washington, DC 20052, USA (e-mail: judywang@gwu.edu).

story of this family, the Schischa family (English version: “Shisha”), is the topic of an almost finished book by Carol G. Vogel, an award-winning author of more than 20 books for children. So there is a lot known about my mother’s family. But, I know almost nothing about my father’s family, except that he was born in New York City (NYC) in 1907, met my mother in Budapest in 1932 when she was 18, fell in love with her, changed his German name to the Hungarian “Székely,” married my mother just before World War II (WWII) and then stayed and lived in Budapest (another option would have been NYC; because of WWII this would have been a better one – retrospectively). My father loved all kinds of glass products, so not surprisingly this became his profession. In the glass factory, where he was one of the leading figures, they produced and exported all kinds of glass products like rear-view mirrors for Volkswagen cars and glasses for the Metropolitan Opera. My father died in 1975. I visited NYC for the first time exactly 10 years after his death. My mother died in 2006, in Budapest, of course. Figure 1 shows Gábor about 30 years apart.

3. STOCHASTIC PATH TO MATHEMATICS

GPW: Could you tell us how you got interested or ended up in the mathematical sciences?

GS: The famous mathematician, Paul Erdős, visited our family when I was about 8 years old and brought messages from family members who lived abroad. At that time, very few people could cross the Iron Curtain—Erdős was one of them. Erdős asked me if I like math. I said, no, it is too boring (my parents were not very proud of me for my honest answer). Then Erdős explained to me how Gauss, when he was about the same age as me at that time, quickly found the sum of the first 100 natural numbers during his class when the teacher wanted them to work silently on this problem and expecting them to take a long time to obtain the answer. As the well-known story goes, the little Gauss noticed that $1 + 2 + \dots + 100 = (1 + 100) + (2 + 99) + \dots + (50 + 51) = 50 \times 101 = 5050$. I was shocked to see this clever trick and started to discover related tricks (all of them, of course, turned out to be well known already, such as the sum of geometric series, etc.). I started to love integer numbers. In elementary school, I was meditating on the question why 7 is a mystical number. I came up with the observation that the series of mystical numbers (lucky or unlucky): 3, 7, 13 are always bigger by exactly 1 than the least common multiples of the first positive integers: $\{1, 2\}$, $\{1, 2, 3\}$, $\{1, 2, 3, 4\}$. I immensely enjoyed the Pythagorean music theory because it explained “harmony” via ratios of positive integers.

GPW: Did you do anything else for fun as a child, besides studying mathematics and discovering mathematical tricks?

GS: In my family, there were several famous musicians and I did not want to “compete” with them. It would have been hard since one of them, Joseph Joachim, a friend and close collaborator of Johannes Brahms, is regarded as one of the most significant violinists of the 19th century, while a sister of my grandmother was a concert pianist. Although I do not play any musical instrument, I was a board member of the Toledo Opera between 2000 and 2006. I love music but, instead of playing music, I started playing water polo. At that time, the Hungarian water-polo team was the World Champion. When I was about 14, it turned out that I could not make the Hungarian National Team of my age group, and because of this, I stopped playing water polo, including swimming, and started to focus on math.

EP: (*interrupting and smiling*) So, one could say then that, in a sense, it was a win-win-win situation for Hungary, you and the world: the Hungarian national team did not get a bad player, You became a mathematician, and the world got a good mathematician!

GS: (*continuing after the interruption*) In my high school years, I wrote essays on the similarity between the axiomatization of geometry in ancient Greece and the “axioms” of social life in the Bible, the Ten Commandments. It was comforting for me to explain the world via simple numbers. Numerals revolutionized our civilization: they expressed abstract thoughts, after all, “two” does not exist in nature, only two fingers, two people, two sheep, two apples, two oranges. After this abstraction, we could not tell from the numerals what the objects were; seeing the signs of 1, 2, 3, ... we could not see or smell oranges, apples, etc. but we could do comparisons, we could do accounting, “statistics,” “statistical inference.” Exact science, statistics, and thus the intellectual world as we know it, was created about 6000 years ago in Mesopotamia through the introduction of number names. As L. Kronecker said in German: “Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk” (in English: “God made the integers, all else is the work of man”). In high school, I was among the best 12 students who were invited to prepare for the high school International Mathematics Olympiad (IMO). In these coaching classes, I made good friends, such as László Lovász, but I never made the top 6, so I never made the IMO team—I was not fast enough in solving problems. Besides mathematics, my hobby was Latin and ancient Greek. In high school, I did not learn English, French or German because we had more teachers who knew Latin and ancient Greek. So, “by accident” I became an “expert” in ancient history from the Biblical times until the Roman period. My other favorite era is 1900 Europe where before and after WWI Budapest played a leading role that resulted in outstanding high schools where students like John von Neumann, Edward Teller, George Pólya and Paul Erdős studied between the two world wars. Except for von Neumann, I met



FIG. 1. Gábor J. Székely. Then (circa 1987) and Now (circa 2017).

all of them. When I met Teller in Budapest, I asked him what was the most important thing he learned from his Ph.D. advisor, the Nobel Prize recipient, Werner Heisenberg. Teller's prompt reply: "Whatever I told him he always interrupted and asked 'Wo ist der Witz?' (in English: "What is the point?"). So, eventually, I always needed to start with the point."

GPW: Was there any particular incident that got you interested in probability and statistics?

GS: Frankly, for a long time probability, not to speak of statistics, was not my favorite area of interest. The following example, and my professor, Alfréd Rényi, changed my mind. Suppose we have 365 students in a lecture hall. Then it is possible that all of them have different birthdays; we need at least 366 students to be 100% sure that two of them have the same birthdays (forget about leap years for now). But, if instead of 100% we can accept 99%, then 55 students are enough to claim that at least two of them have the same birthdays. How come 1% "flexibility" can cause such a dramatic change from 366 to 55? It should mean that randomness/probability must be a very powerful tool in engineering, etc. Even if we want more certainty, like 99.9% we need just 68 students instead of 366. The difference is huge and this is relevant to statistics, too. A tiny little extra risk can change the world!

4. BUDAPEST YEARS: RÉNYI, KOLMOGOROV, ERDŐS, LINNIK

My statistics and probability research started in Budapest, Hungary. In 1968–1970, I was a student of Alfréd Rényi (see Figure 2). I enjoyed every minute of his lectures and also our conversations on philosophy, ancient Greek history and axiomatization of science, but before I could finish my graduate studies, Rényi suddenly died of

lung cancer in 1970. Some of my other outstanding professors, like Imre Csiszár and Paul Révész, were teaching information theory and ergodic theory, but these areas were not too close to my heart at that time. So, in 1970 I was left without Ph.D. advisors, but I could start to teach at Rényi's former department. In 1972, Budapest organized the European Meeting of Statisticians and one of my heroes, A. N. Kolmogorov, was a distinguished participant. I asked him if he would accept me as a student for the Candidate Degree. He handed over an interesting problem to me and said: yes, if you can solve this. Our conversation lasted for a few seconds only but after that I was working on his problem for 3 years, solved the problem (on the limiting behavior of permanents of random variables of $m \times n$ matrices as $m \leq n$, $n, m \rightarrow \infty$, $m/n \rightarrow c$ where $0 \leq c \leq 1$; the special case $m = 1$ is the classical partial sum) and defended my Candidate Degree in 1976. For initial and related results, see [7] and [10].

After defending my Candidate Degree, I started to write a book on probability and statistics paradoxes. The Hun-

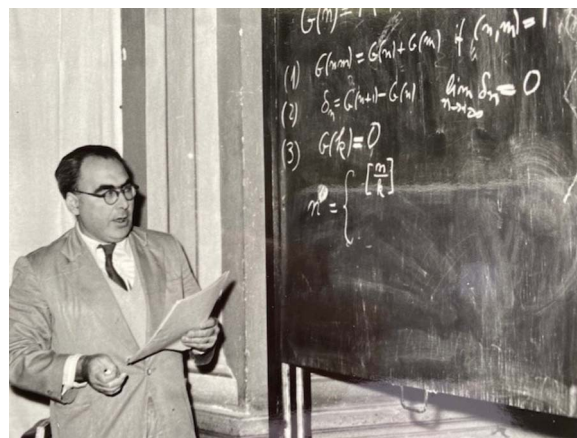


FIG. 2. Gábor's Ph.D. adviser, Alfréd Rényi (circa 1965).

garian original came out in 1982. Later it was published in English by Reidel–Kluwer [19]. Then Springer published a paperback version without my permission and when I called their attention to this “minor problem” they deleted my book from their web pages, which is probably the worst remedy! The paradox book was also translated to both Russian and German.

In the 1970s, many talented probabilists and statisticians in Budapest were focusing on the so-called Komlós–Major–Tusnády (KMT) approximation, which was considered the most important discovery in probability and statistics in the post-Rényi era in Budapest. For more information, you can consult the book by Csörgő and Révész [3]. I could not contribute anything to this topic. On top of that, for a long time I could not travel or study abroad because of the “sin” that my father was born in NYC, thus by definition I was considered a “capitalist.” My chances to meet foreign scientists were limited to Paul Erdős’ conferences and other typical “Hungarian math conferences” of the Bolyai Society. This is where I met Yuri Linnik and Bruno de Finetti. In 1969, Linnik happened to sit next to me at a conference dinner table, I introduced myself and asked him why in his publications he calls two-point distributions “prime” when in fact in the number theoretic sense they are “irreducible” (in the number theoretic sense, prime distributions are nondegenerate and can divide a convolution product of probability distributions only if they divide at least one of the factors, while irreducible distributions cannot be the convolutions of two nondegenerate distributions, which is clearly true for all two-point distributions). Linnik’s short answer was: “Well?!” After this “encouragement” I started to work on the irreducible-prime problem with Imre Z. Ruzsa, a brilliant young number theorist. We proved that, in fact, in the convolution semigroup of probability distributions on the real line, there are no prime distributions at all; see [17]. As a consequence, there is no confusion if we call irreducible distributions prime, at least on the real line. In a sense, this answered my question to Linnik. Jointly with Ruzsa, we wrote a book, *Algebraic Probability Theory*, on this topic [18]. In 1988, we were awarded the Rollo Davidson Prize of Cambridge University for our work on algebraization of probability theory. As a by-product of this algebraization, we could answer a fundamental question on negative probabilities (see the next section).

YG: Could you tell us a bit on how mathematics and statistics research was organized in Hungary during those times? Also, collaborations, travel, funding, work with students, etc., and challenges and opportunities?

GS: Mathematical and statistical research was very much different from the research in the U.S. At that time, even today, almost all faculty members and students were Hungarians, and all my collaborators were Hungarians. It was a real culture shock to me in the U.S. to collaborate with Russians, Italians, Chinese, etc. who were faculty and students in the same department.

5. ON NEGATIVE PROBABILITIES

YG: Could you tell us about the surprising notion of negative probabilities?

GS: In the Kolmogorov system of probability theory, probabilities are nonnegative numbers in the interval $[0, 1]$. So, case closed: negative probabilities do not make any sense. After all, what would be the meaning of negative frequencies? But, of course, $-\$3 + \$5 = \$2$. Can we do something similar with probability distributions? Is it true that if P is a signed probability distribution such that the probability of the whole space is 1, then we can always find two nonsigned probability distributions, Q and R , such that the convolution of P and Q is R . The answer is: YES (see [16]) and this can be considered a fundamental result for negative probabilities. Directly, we cannot observe a signed probability distribution P that has events with negative probabilities, but if we add a suitable independent error with nonsigned probability distribution Q , then we always get a nonsigned, observable, probability distribution R , which is the convolution of P and Q .

An example for negative probabilities is the half coin. I introduced this strange “pseudo-coin” in [20]. It is a signed probability distribution with generating function $p(z)$ such that its square, $p^2(z) = (1+z)/2$ is the generating function of a fair coin; $p(z) = \sum_{k=0}^{\infty} p_k z^k$ where p_k is the signed probability of taking the value k . One can easily compute that $p_k < 0$ for $k = 2, 4, \dots$ (If we substitute $z = 1$, then we can see that the sum of the signed probabilities $\{p_k, k = 0, 1, 2, \dots\}$ is 1; but if we substitute $z = -1$, then we can see that the sum of the absolute values of the signed probabilities $\{|p_k|, k = 0, 1, 2, \dots\}$ is $\sqrt{2}$.) According to our fundamental theorem above, we can always find two (nonsigned) generating functions $q(z)$ and $r(z)$ such that $p(z)q(z) = r(z)$. Recently, negative probabilities, and also negative energies, have become crucial in modern physics of elementary particles. This approach started with two Nobel Prize recipients, Eugene Wigner and Paul Dirac. Espen Gaarder Haug is a pioneer in applying negative probabilities in finance (see his publications). He started his career as a guard of the Norwegian King, then became a derivative trader for J. P. Morgan where we became friends in 2003 when we both “exercised the option to meet in New York City.”

6. ARE WE ALL NORMAL?

JW: What other statistical problems fascinated you during these times?

GS: According to a classical result of C. F. Gauss, normal distributions can be characterized by the property that the maximum likelihood estimator of their location parameter is the arithmetic mean of the observations. But what if we replace the arithmetic mean by general L-statistics of observations (these are linear functions of order statistics)? In a paper [2] with Z. Buczolic, we proved

that apart from the arithmetic mean and the corresponding normal distribution, L can be a maximum likelihood estimator of the location only if at most two coefficients in L are non-zero, such as the midrange for uniform distributions or the minimum for exponential distributions.

7. OBERWOLFACH RESEARCH INSTITUTE

YG: Prior to your going to the United States, you visited Oberwolfach in Germany and got connected with German researchers, to the extent that you mentioned the possibility of doing research in Germany post-retirement. Please tell us about your interactions with German colleagues.

GS: In the 1980s, I gave several talks in Oberwolfach where I made many friends. After my Oberwolfach conference talk on Gauss and L -statistics, I got acquainted with Ursula Gathers. She was then an energetic young professor from Dortmund, Germany. We were chatting about the importance of initiating a new journal of rejected papers because too many revolutionary papers are rejected: they are “too risky” to be published. According to our plans, when one submits a paper for publication to the *Journal of Rejected Papers* one needs to accompany at least three letters of rejections from previous submissions. We never had time to create such a journal since Ursula quickly became the President (Rector) of the Technical University of Dortmund.

GPW: Do you think such a journal should be pursued at this time?

GS: Yes, because in science, and also in statistical science, not like in politics, it is good to embrace extremes and this is not what typical cautious editors do who are afraid of publishing something incorrect at the expense of rejecting revolutionary ideas. Unfortunately, this “cautiousness” refers not only to scientific journals. For example, is NSF too cautious? I am, of course, not objective. Ask those whose proposals were rejected.

Other German statisticians I have regular joint research with include Norbert Henze, Prorector of Karlsruhe University until 2009.

8. VISITING THE UNITED STATES

GPW: Tell us about your first visit to the U.S.

GS: In 1985, on my first visit to the U.S., I was almost 40. I tried to meet as many of my “heroes” as I could and give talks at their universities. Examples include Herbert Robbins (at that time he was the Chair of the Department of Statistics at Columbia University) and George Pólya (Stanford University). Robbins asked me to translate his classical book, jointly written with Richard Courant, “What is Mathematics?” into Hungarian. I did and submitted the manuscript for publication; it turned out that the Hungarian translation of the book was already published many years earlier. Hmm...

My meeting with Pólya was also interesting. I phoned him and asked him which language he prefers for our conversation: English, Hungarian or German? He replied: ancient Greek. I am not sure if he was joking, but we continued in English. He told me that his most important advice to young researchers is engraved into the walls of a building at Stanford: “If you want to give a good talk, you need two ingredients: (i) you need to have something to say and (ii) if you happen to have two different things to say, then first finish the first one before you start the second one.” As I can recall, I found this quotation in 1985 engraved into a wall of Stanford, but at my later visits at Stanford, I could not find it again.

9. BUDAPEST INSTITUTE OF TECHNOLOGY AND THE BGSU YEARS

YG: Could you tell us how you ended going to Bowling Green State University?

GS: The story goes back to pre-WWII Vienna, Austria where Eugene Lukacs, the author of the classic textbook *Characteristic Functions*, grew up. When Germany annexed Austria in 1938, many Jewish mathematicians, including Lukacs, emigrated to the USA. See Figure 3 for a picture of Lukacs with Rényi. Here, he renewed his acquaintance with Abraham Wald whom he had met many times in Vienna. Wald influenced Lukacs to become interested in probability and statistics, the latter still almost unknown in Central Europe during that time. In 1955, Lukacs joined the Catholic University of America where he organized the Statistical Laboratory and became its first and only director. When Lukacs retired in 1972 the Department of Mathematics at Bowling Green State University (BGSU) invited him and two of his students who were his colleagues at CUA, Radha Laha and Vijay Rohatgi, to initiate and organize the Ph.D. program there. A few months after the death of Lukacs in 1987, Rohatgi invited me to BGSU for one semester in order to continue the work of Lukacs, and I availed of this opportunity in



FIG. 3. Alfréd Rényi and Eugene Lukacs (circa 1960).

1988, where incidentally I first met you, Edsel, since you were a young assistant professor then at BGSU.

After spending one semester at BGSU in 1988, and then another one at Yale in 1989 (where I became friends with Grace Wahba who also visited Yale), I was invited back to Budapest to organize a Department of Stochastics at the Budapest Institute of Technology (Technical University of Budapest). I was very happy to work on this project. This department still exists today! It was at the Budapest Institute of Technology where I got acquainted with two outstanding statisticians, László Györfi and Gábor Lugosi; in the mean time, László has retired but is still a very active researcher in Hungary, while Gábor moved to Barcelona where he is a research professor at Pompeu Fabra University.

Between 1990 and 1991, I was invited back to Bowling Green to be appointed as the first Eugene Lukacs Distinguished Professor at BGSU. This was a visiting professorship supported and funded by the State of Ohio. The Lukacs Distinguished Professor, aside from giving regular lectures and seminars, spearheaded the organizing of the annual Lukacs Symposium on a topic of his/her choice. In 1995, I was then offered, and I accepted a tenured full professorship at BGSU. I stayed at BGSU as professor for 11 years prior to joining the National Science Foundation. I had six outstanding Ph.D. students there, including Maria L. Rizzo and G. Jay Kerns. During my tenure at BGSU, I was also able to invite several outstanding statisticians and probabilists for the Lukacs Distinguished Visiting Professorship, such as Bradley Efron, Peter Bickel, C. R. Rao (see Figure 4), N. Henze, E. Seneta, I. Karatzas, etc.

10. BRUNO DE FINETTI AND EXCHANGEABILITY

JW: You have papers that span classical statistics and Bayesian statistics. How did you get enamored with the Bayesian approach and what would you consider as your main contribution to Bayesian statistics?

GS: I met de Finetti in 1969 at one of Erdős' conferences. We discussed de Finetti's famous representation theorem: an infinite sequence of exchangeable random variables is always a sequence of conditionally independent, identically distributed random variables given a suitable random variable as a condition. This is important in Bayesian statistics. Unfortunately, de Finetti's theorem is not always true for finite exchangeable sequences. As I proved many years later, jointly with one of my Ph.D. students, J. Kerns [8], the remedy for finite sequences is conditioning on random variables that can take values with negative probabilities (they can have signed probability densities – or negative probabilities; see [20]).

11. JOYS OF POSING PROBLEMS

EP: Gábor, I recall you have a penchant for posing mathematical problems. When I was a young colleague of yours at BGSU in the 1990s, I recall you coming to my office and asking: “Edsel, could you give me a nontrivial ancillary statistic if X_1, X_2, \dots, X_n is a random sample from the Bernoulli distribution with unknown parameter $\theta \in (0, 1)$?” I thought immediately that the problem was trivial, since under random sampling from a continuous distribution, the rank vector is always a nontrivial ancillary statistic, but this turned out to be not-so-trivial after all. We ended up writing a neat paper (see [13]) on this problem which established that under discrete distributions allowing a complete sufficient statistic with an atom, there exists no nontrivial ancillary statistic! So, tell us more about your penchant for posing mathematical problems, and do you have some current open problems?

GS: There are mathematicians or scientists in general who are good at solving open problems of others. This is not me. I wish I were a better problem solver. My strength is to pose important and interesting new questions. Almost all of my joint papers and joint research started with my questions others found interesting or important. Both posing good questions and solving important questions of

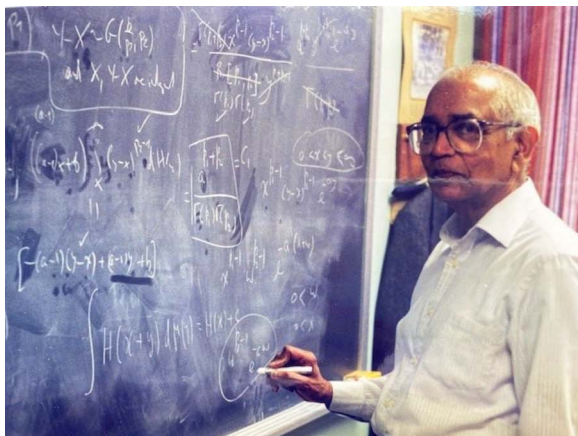


FIG. 4. *Lukacs Professorship at BGSU with C. R. Rao, 1999.*

others are essential. I don't really know why posing problems is my strength. Perhaps, it is the influence of Paul Erdős?

My "simplest" open problems these days are as follows:

(i) If we want to generalize the t -test from Gaussian distributions to all symmetric distributions, then we can reduce this problem to the following simple question; see [21]: in the n -dimensional Euclidean space find the maximum number of vertices of a given cube we can cover by a sphere of a given radius r . For some special cases, for example when n is small like $n = 2$ or $n = 3$ this is a high school exercise but surprisingly, the general solution (especially when it becomes useful for statistical tests) is open.

(ii) There is another simple looking problem that was unsolved for decades: how many tickets you need to buy in a 90 choose 5 lottery to be sure that you have 2 hits on at least one ticket. Just before I left Hungary and moved to the U.S., I could solve this problem with two other Hungarian mathematicians: the surprising and nice answer is exactly 100 (see [6]). Nobody knows the answer for 3 hits.

(iii) According to a classical theorem of Dirichlet, if the positive integers a and d are coprime, then in the arithmetic progression $S_n = a + nd$, $n = 1, 2, \dots$ there are infinitely many primes. But what if the differences d are random and we consider $S_n = X_1 + X_2 + \dots + X_n$, $n = 1, 2, \dots$ as a random arithmetic progression when the terms X_1, X_2, \dots are i.i.d. and take positive integers only? Suppose that the least common divisor of all numbers the i.i.d. random variables X_n take with positive probabilities is 1 and denote the m th prime number by p_m . By the Borel–Cantelli lemma, if $\sum_{n,m} P(S_n = p_m) < \infty$, then S_n and p_m meets (equal) only finitely many times with probability 1. But what if this Borel–Cantelli sum is infinite? My conjecture is that then they will meet infinitely often with probability 1. This is, however, an open problem. What is known is published in my paper [15] with I. Z. Ruzsa.

For the solution of another open problem (Fermat's last theorem for rational exponents), I needed to wait for American collaborators, Curtis Bennett and Andrew Glass, who were then colleagues at BSGU; see [1].

12. LIFE IN BOWLING GREEN, OHIO

GPW: What was your biggest challenge after moving to Bowling Green State University in Ohio?

GS: The biggest challenge was the small town nature of Bowling Green, Ohio. Before Bowling Green, I never lived in a city with less than a million people. But I really enjoyed to learn this new lifestyle, meaning that I knew almost everybody there from the Campus or from restaurants, etc. An equally big challenge was that by 1995,

when I moved permanently to the U.S., my children were of college age and they stayed in Europe (although my son spent a few semesters in Bowling Green).

First of all, I wanted to move to the U.S. for three reasons:

(i) personal: my father was born in New York City so understandably I have always had some affections to NYC and to the U.S.;

(ii) political (this does not need much explanation);

(iii) scientific: U.S. universities are fantastic scientific melting pots.

After moving to BGSU, the biggest challenge for me was to harmonize my childhood dreams about America with the reality of America. An American small town is very much different from a vibrant European capital. On the other hand, the scientific side of my dreams were in complete harmony with my dreams: in Hungary all my collaborators were Hungarians, but in the U.S. they were Russians, Germans, Americans, Italians, Chinese, Indians, Turkish, Filipinos, etc.

13. ACADEMIC ADVISOR OF MORGAN STANLEY

GPW: We heard that you were paid highly by Morgan Stanley! Can you tell us this story and how it all started?

GS: Morgan Stanley paid good money, and I badly needed that because I lost a lot of money on all my houses; I always moved at the wrong time, for example, in 2008 when I moved to Washington, D.C. This was the time of global financial crisis and recession, the real estate market collapsed and my house in Ohio lost most of its value. But let us start at the beginning. In 1995, I was offered a tenured professorship at Bowling Green State University. I was asked to teach actuarial science and be responsible for the actuarial science program there. As I was told, actuarial majors bring a lot of money to the university. Unfortunately, I had no idea about actuarial science except one sentence I learned in my Scientific Socialism undergraduate course—Hungary was then in the orbit of the Soviet Union. According to Scientific Socialism by the time I will be as old as I am now, there will be communism everywhere on Earth, and hence there will be no money at all. But this one sentence did not seem to be enough for a three-semester course, so I started to learn actuarial science and financial mathematics from outstanding texts. I became familiar with the topic and in 1999 I could even predict the dot-com collapse of the stock market in 2000. In these financial predictions, I could use "negative probabilities" that I introduced. Both Morgan Stanley and J. P. Morgan approached me to explain my ideas. They heard about my lectures on this topic in Finland, but mainstream math journals refused to publish my results. Their main concern was that nobody would believe the financial applications, for example, the collapse of the



FIG. 5. Gábor with U.S. Senator from Connecticut, Joseph Lieberman, 2009.

financial market. The financial world was more flexible. They heard about my prediction of the dot-com bubble and soon after the bubble I became an academic advisor of Morgan Stanley in New York, and then academic advisor of Bunge in Chicago. As they said, they do not care about the philosophy of negative probabilities as long as they can make money from them. The financial magazine, *Wilmott*, published my results, not the Mathematical Monthly. After that, I helped to establish the Morgan Stanley Mathematical Modeling Centre in Budapest (2005) and the Bunge Mathematical Institute (BMI) in Warsaw (2006) to provide quantitative analysis to support the firms' global business. At the opening ceremony of the Bunge Mathematics Institute, I could shake hands with the president of Poland, Lech Kaczynski, who later died in a plane crash in 2010. (I never came close to any president of Hungary nor to any president of the U.S., though I came quite close to U.S. Senator Joe Lieberman; see Figure 5.) The Morgan Stanley Budapest office is still very active and is located at the Danube in the Millennium City Center.

14. STUDENT MENTORING

YG: Did you mentor many students and any advice or insights on mentoring students?

GS: I did not have many Ph.D. students but the ones I had were excellent. Here is the story of one of them. Right after I was parachuted to Bowling Green from Europe, all my esteemed colleagues at BGSU told me: if you want good graduate students, choose Chinese ones. But, by the time I got the list of graduate students, one single Chinese name was left: Maria Hong. So I picked her, but to my shock, when she showed up in my office, she did not look Chinese. She told me that her husband is Chinese, but if I accept her as a Ph.D. student, then she will change her name back to her maiden name, Maria Rizzo, so that she would not mislead anybody else. She turned out to be my best student in the U.S. My advice: *It is better to have one outstanding student, than ten mediocre ones.*

15. UNCERTAINTY PRINCIPLE OF GAMES

EP: Your knowledge of physics have been highly beneficial in your statistical research to the extent that this is one of the reasons for your getting elected as Fellow of the Institute of Mathematical Statistics. Could you tell us an example of where you could apply a physics concept in statistics?

GS: Already in the 1980's I started to focus on applying physics principles into statistics. In 1927–1928, two fundamental papers changed our views on matter and mind. Werner Heisenberg's Uncertainty Principle in quantum mechanics, published in 1927, led to a probabilistic description of matter (unlike Newton's deterministic laws), and John von Neumann's Minimax Theorem of Zero-Sum games, published in 1928, led to a probabilistic description of our mind. I tried to combine the two and apply the uncertainty principle to minimax strategies that are very important in statistics. For simplicity, suppose that two players have finitely many strategies: n and m , respectively, and the payoff matrix is $(a_{i,j}, i = 1, 2, \dots, n, j = 1, 2, \dots, m)$. Without loss of generality, we can suppose that the smallest $a_{i,j}$ is 0 and the biggest one is 1. Denote the minimax strategies of the two players by P and Q ; these are probability distributions on the possible n or m pure strategies. According to our uncertainty principle of game theory, the entropy of P (and the entropy of Q) cannot be smaller than the entropy $H(h, 1-h)$ of a two-point distribution h and $1-h$, where $h := \min_j \max_i a_{ij} - \max_i \min_j a_{i,j}$, the so-called commutator of the two (nonlinear) operators \min and \max . The lower bound, $H(h, 1-h)$, is sharp. Based on this uncertainty principle, we can approximate the optimal (but typically complex) minimax strategies supported on many points by much simpler mixed strategies supported on two points only. This is very important in military and other applications, where too many options are too confusing. For details, see my paper with Maria Rizzo [23]. I discussed this topic many times with the brilliant Larry Shepp.

16. ON ENERGY OF DATA

GPW: One of the most influential concepts you introduced in Statistics is the notion of the Energy of Data. Could you please expound on this notion and how it came to be?

GS: Another extremely important notion of physics is energy. Newton's potential energy is a function of distances between physical objects. Based on this idea, I introduced a general notion of energy for data in metric spaces. For simplicity, suppose that all observations (all data) are coming from a given Hilbert space where the distance between points is denoted by d . Then the energy distance between independent random variables X

and Y (or of their probability distributions) is defined as the square root of the nonnegative quantity:

$$D^2(X, Y) := 2Ed(X, Y) - Ed(X, X') - Ed(Y, Y'),$$

where X' is an i.i.d. copy of X and Y' is an i.i.d. copy of Y and we suppose the underlying expectations denoted by E are always finite. One can show that D is always a metric. If instead of random variables we have statistical data of size n , m , we simply replace the expectations E by arithmetic means:

$$D^2 := \frac{2}{nm} \sum_{i=1}^n \sum_{j=1}^m d(X_i, Y_j) - \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d(X_i, X_j) - \frac{1}{m^2} \sum_{i=1}^m \sum_{j=1}^m d(Y_i, Y_j).$$

The source of data energy is D . This can also be viewed as energy in physics if we multiply D^2 by a constant force of magnitude F ; that is, energy $\mathcal{E} = F \cdot D^2$. The signs in D^2 resemble the computation of electrostatic potential where the signs depend on the charges. For statistical inference, we do not need to know F . We can simply take $F = 1$, but F can also be viewed from the physics perspective as a “God-given” constant like the speed of light or the gravitational constant. This F would play an important role if we wanted to “free” the data energy as a counterpart of nuclear energy. We can also take the power $0 < \alpha < 2$ of all distances in the definition of energy and it remains nonnegative and equals 0 if and only if X and Y are identically distributed. This α -energy is a generalization (to two random variables) of the classical Riesz energy $E|X - X'|^\alpha$ in Euclidean spaces (where α is a real number). The energy distance of data became very widely applicable in modern statistics because instead of working with very complex data we can always focus on their real valued distances in metric spaces; see [22, 25, 27]. So, after our long journey we are back to real numbers. In the energy world of statistics, we can forget vectors, matrices, functions and more complex data, it is always sufficient to deal with their real valued distances, with real numbers, as long as the data are in metric spaces. Applications (see [14, 25]) include testing for normality, hierarchical clustering, analysis of variance, change-point detection, etc. In the classical area of testing multivariate normality the best modern consistent test, the BHEP test, is due to Norbert Henze (Karlsruhe, Germany) and to his collaborators. Our energy test [12] based on energy distance became a very powerful competitor. Although I have no joint papers with Henze, we met many times and we always had very interesting “duels.” There are many contributors to the applications of energy distance, among others: N. K. Bakirov (who was killed in a car accident in Ufa, Russia), M. L. Rizzo, L. B. Klebanov, R. Lyons, R. Davis, J. Fan, X. Huo, B. Sen, R. Tibshirani, G. Tilmann, D. Richards,

D. Edelman, T. F. Móri and D. Matteson. Distance correlation is probably the most important application. On kinetic energy of data, see [25, 27].

17. ON DISTANCE CORRELATION

EP: Speaking about distance correlation, your most cited paper is on this topic with Rizzo and Bakirov. Can you tell us a bit more how the idea came about? What are the advantages of this correlation measure over other measures of associations, such as the well-known Pearson’s correlation coefficient?

GS: The paper on distance correlation in the *Annals of Statistics* is my most cited paper [28], the number of citations is more than 2000. Distance correlation is defined with the help of distance covariance the same way as Pearson’s classical correlation is defined via classical covariance. Michael Newton writes about distance covariance in introducing my 2009 discussion paper [24] in the *Annals of Applied Statistics*:

I recall a great sense of excitement in the seminar room in Madison after Professor Székely presented the astonishing findings about distance covariance, in the spring of 2008. It was one of the best statistics seminars I could remember. Since before computers statisticians have held up Pearson’s correlation coefficient as the most essential measure of association between quantitative variables. R. A. Fisher’s reputation was sealed, in part, by solving the distribution of this statistic, and so much of linear-model methodology relates to it. And all the time we’ve had to add the caveat about independence following zero correlation only if the data are jointly normal. Spearman’s rank correlation has substantial practical utility in cases where normality is unreliable, but the goal to have a bona fide dependence measure seemed to have been beyond the scope of ordinary applied statistics. Some valid measures did exist, but being driven by empirical characteristic functions, they were too complicated to enter the toolkit of the applied statistician. Distance covariance not only provides a bona fide dependence measure, but it does so with a simplicity to satisfy Don Geman’s elevator test (i.e., a method must be sufficiently simple that it can be explained to a colleague in the time it takes to go between floors on an elevator!). You take all pairwise distances between sample values of one variable and do the same for the second variable. Then center the resulting distance matrices (so each has column and

row means equal to zero) and average the entries of the matrix, which holds component-wise products of the two centered distance matrices. That's distance covariance between the two variables. The population quantity equals zero if and only if the variables are independent, whatever be the underlying distributions and whatever be the dimension of the two variables. The depth of the finding, the simplicity of the statistic and the central role of statistical dependence make this an important story for our discipline.

The population distance covariance is simply the energy distance between the joint distribution of X, Y and the product of their marginals. But why exactly do we need a new measure of dependence? We have so many other measures for this? The most applied one is Pearson's correlation, but it has two major disadvantages: (i) it can be 0 even if the variables are not independent and (ii) it is defined for real-valued variables/data only, not for general metric space-valued variables. Distance correlation does not have these disadvantages. But then there is maximal correlation or the maximum information coefficient. They do not have these negative properties, but they are not continuous with respect to weak convergence of distributions, so it can happen that the empirical measure of maximal correlation and of maximal information coefficient is close to 1, while the theoretical measure of dependence is 0. Distance correlation is a remedy to this problem, too. Distance correlation is continuous with respect to weak convergence of probability distributions; see [11] and [22]. Distance correlation is easy to compute, it is almost as easy to compute as the classical correlation, while maximal correlation and maximal information coefficient are absolutely nontrivial maximizations. On the generalization of distance correlation to metric spaces, see Lyons [9]. On the connection between distance covariance and Brownian motion, see [24]. Distance correlation can be applied to "dismantle" Mantel's test and enables the introduction of an improved version (see [26]). On the connection between Pearson's classical correlation and distance correlation, see [5].

18. SOJOURN AND TENURE AT THE NATIONAL SCIENCE FOUNDATION

JW: Gábor, you have been a mainstay, as a Program Director, in the Statistics Program of the Division of Mathematical Sciences at NSF since 2006 until your retirement in February 2022. Could you tell us how you ended up at NSF as a Program Director, and regarding your eventual decision to become a permanent program director in the Statistics Program?

GS: My NSF life started with an invitation to be a panelist at an NSF statistics review panel. The invitation came

from Grace Yang. I liked the atmosphere of the panel and the program directors liked my reviews. Soon after that, in 2006, I was invited for a 2-year program director position in statistics and probability. I accepted this offer. Within a year, all of my permanent program director colleagues in probability and statistics accepted other job offers so I was left with two options: (i) go back to teach at universities and leave the NSF statistics and probability boat to sink or (ii) stay at NSF and save the statistics and probability program. I stayed. Since 2006, I have been working at the National Science Foundation as a permanent Program Director. In 2008, I was asked to choose between the Probability Program and the Statistics Program, as I could not be the program director of both. I chose statistics for two reasons. Statistics is a huge area with many important grant applications and the changing world of statistics from mathematical sciences toward data science was in the air. Also, my personal research interest shifted toward statistics, especially toward applications of physics concepts, like energy, to statistics. I am proud of the fact that the statistics program could support all major important areas of statistics in a fast moving world where statistics is not the same as mathematical statistics. Modern statistics combine math, computer science, artificial intelligence, etc. The empirical nature of statistics can be compared to empirical physics. It can easily happen that very soon grant applicants will be robots with artificial intelligence, and will be judged by robots with artificial intelligence in an artificial panel. But by then I will do something else, not in the NSF building. I decided to retire in 2022.

YG: What are the most surprising trends you have noticed in mathematics and statistics, either successes or failures? What trends do you expect in the future?

GS: The trends are nicely shown by the names of the statistics journals of the Institute of Mathematical Statistics (IMS). Until 1972, the top statistics journal was the *Annals of Mathematical Statistics*. In 1973, it was split to *Annals of Probability* and *Annals of Statistics*, so probability theory was somewhat distanced from statistics, the latter having become less mathematical. Then in 2007 a new journal was created: *Annals of Applied Statistics*. Nowadays statistics has become a pillar of Data Science and thus involves machine learning, computer science, artificial intelligence (AI), etc. We need to keep a good balance between experimental, applied and theoretical statistics. I would not be surprised to see grant proposals in the near future written by AI and reviewed by AI panelists. But this is beyond my "old style" world view.

YG: Can you please give an advice to aspiring new and not so new principal investigators in terms of their proposal writing?

GS: According to the Roman poet, Juvenal, it is difficult not to write satire. It is even more difficult to give advice on how to write grant proposals. The good news

is that NSF has always been open to any and all important new ideas, but the grant proposal should make these ideas accessible to a wider audience. The wider the better. It does not matter if your university is not in the top tier as long as your new ideas are.

YG: Gábor, I am curious what your views are regarding “birds and frogs” in the statistical sciences in light of Freeman Dyson’s Einstein Lecture titled “Birds and Frogs” (see [4])? In this lecture, Dyson, an English-American theoretical physicist and mathematician, begins with the observation:

Some mathematicians are birds, others are frogs. Birds fly high in the air and survey broad vistas of mathematics out to the far horizon. They delight in concepts that unify our thinking and bring together diverse problems from different parts of the landscape. Frogs live in the mud below and see only the flowers that grow nearby. They delight in the details of particular objects, and they solve problems one at a time.

— Freeman Dyson in his Einstein Lecture

GS: My answer is that there are many animals and we need all of them. In terms of Dyson’s characterization, I am surely not a frog, but I cannot always fly like a bird!

19. ON FAMILY AND RETIREMENT

GPW: Now that you are officially retired from NSF starting the beginning of March 2022, what are your thoughts about family?

GS: I have a nice family: wife Judit (see Figure 6), two children Szilvia and Tamás (see Figure 7), and six grandchildren: Elisa, Anna, Michaël who live in Brussels, Belgium and Lea, Esther, Avi who live in Basel, Switzerland (see Figure 8). Three of my grandchildren already show interest in math. Others in music. Will they be statisticians



FIG. 6. Gábor and his wife, Judit, at the Residency of the U.S. Ambassador to Hungary, 2017.

or mathematicians or musicians? I don’t know. Even the oldest is only 16 years old. Where should I live after retiring from NSF? The moon is about equal distance to all of them!

GPW: Do you have any projects that you would like to pursue during your retirement?

GS: My next project is the application of data energy to defense industry. Unfortunately, this has become a hot topic. The Energy of Data and Distance Correlation is the title of my next book. According to our plans, it will come out in 2023.

YG: Gábor, my other question of personal interest: what do you think on the hypothesis of relationship between music and mathematics and what is your favorite piece of music? I know there are many, but if you can choose the one?

GS: On the relationship of music and math: I know that Paul Erdős was orthogonal to music, for him even Rachmaninov was noise. I cannot really choose a single favorite piece of music, it depends on my mood but the



FIG. 7. Gábor with his children, Szilvia and Tamás.



FIG. 8. Gábor's grandchildren, 2021.

Russian musical soul is very close to my heart, maybe because they are like birds in Freeman Dyson's Einstein Lecture "Birds and Frogs." I have no idea why Rachmaninov and Tchaikovsky are my favorites. By the way, I wrote a review of the 1989 book by Philip Davis titled *The Thread: A Mathematical Yarn*, which is about the first name of Pafnuty Chebyshev and on the Russian soul; I wish I knew where my notes were published (in Math Reviews or ... ?) Davis' book on the word Pafnuty is a poem; I even suggested a new verb for Webster: "to pafnuty," which means to work with great zeal and affection on something nobody really cares. I guess I was asked to write this review because I translated into Hungarian another book of Davis, joint with Hersh: *The Mathematical Experience*.

GPW: Gábor, it has been a pleasure and honor for all three of us to be your colleagues at NSF. Thank you for agreeing to have this series of conversations with you, virtually via Zoom, via e-mail and in-person. We would like to thank you very much for giving us a window in your very interesting and colorful life and career. Before we close, any final thoughts?

GS: Let me finish with the words of Albert Einstein: "The most beautiful experience we can have is the mysterious. It is the fundamental emotion that stands at the cradle of true art and true science." I hope that in my next projects I shall have as talented and enthusiastic partners with this fundamental emotion as I have had so far.

ACKNOWLEDGMENTS

Yulia R. Gel, Edsel A. Peña and Huixia (Judy) Wang were then Rotating Program Directors in the Division of Mathematical Sciences (DMS) at the National Science Foundation (NSF) when the interviews were conducted. As such, part of this material is also based upon work supported by (while serving at) the National Science Foundation. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Na-

tional Science Foundation and/or the Office of Naval Research.

REFERENCES

- [1] BENNETT, C. D., GLASS, A. M. W. and SZÉKELY, G. J. (2004). Fermat's last theorem for rational exponents. *Amer. Math. Monthly* **111** 322–329. MR2057186 <https://doi.org/10.2307/4145241>
- [2] BUCZOLICH, Z. and SZÉKELY, G. J. (1989). When is a weighted average of ordered sample elements a maximum likelihood estimator of the location parameter? *Adv. in Appl. Math.* **10** 439–456. MR1023943 [https://doi.org/10.1016/0196-8858\(89\)90024-9](https://doi.org/10.1016/0196-8858(89)90024-9)
- [3] CSÖRGŐ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics. Probability and Mathematical Statistics*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York. MR0666546
- [4] DYSON, F. (2009). Birds and frogs. *Notices Amer. Math. Soc.* **56** 212–223. MR2483565
- [5] EDELMANN, D., MÓRI, T. F. and SZÉKELY, G. J. (2021). On relationships between the Pearson and the distance correlation coefficients. *Statist. Probab. Lett.* **169** Paper No. 108960, 6. MR4163211 <https://doi.org/10.1016/j.spl.2020.108960>
- [6] FÜREDI, Z., SZÉKELY, G. J. and ZUBOR, Z. (1996). On the lottery problem. *J. Combin. Des.* **4** 5–10. MR1364095 [https://doi.org/10.1002/\(SICI\)1520-6610\(1996\)4:1<5::AID-JCD2>3.3.CO;2-W](https://doi.org/10.1002/(SICI)1520-6610(1996)4:1<5::AID-JCD2>3.3.CO;2-W)
- [7] HALÁSZ, G. and SZÉKELY, G. J. (1976). On the elementary symmetric polynomials of independent random variables. *Acta Math. Acad. Sci. Hung.* **28** 397–400. MR0423491 <https://doi.org/10.1007/BF01896806>
- [8] KERNS, G. J. and SZÉKELY, G. J. (2006). De Finetti's theorem for abstract finite exchangeable sequences. *J. Theoret. Probab.* **19** 589–608. MR2280511 <https://doi.org/10.1007/s10959-006-0028-z>
- [9] LYONS, R. (2013). Distance covariance in metric spaces. *Ann. Probab.* **41** 3284–3305. MR3127883 <https://doi.org/10.1214/12-AOP803>
- [10] MAJOR, P. (1999). The limit behavior of elementary symmetric polynomials of i.i.d. random variables when their order tends to infinity. *Ann. Probab.* **27** 1980–2010. MR1742897 <https://doi.org/10.1214/aop/1022677557>
- [11] MÓRI, T. F. and SZÉKELY, G. J. (2019). Four simple axioms of dependence measures. *Metrika* **82** 1–16. MR3897521 <https://doi.org/10.1007/s00184-018-0670-3>
- [12] MÓRI, T. F., SZÉKELY, G. J. and RIZZO, M. L. (2021). On energy tests of normality. *J. Statist. Plann. Inference* **213** 1–15. MR4182996 <https://doi.org/10.1016/j.jspi.2020.11.001>
- [13] PEÑA, E. A., ROHATGI, V. K. and SZÉKELY, G. J. (1992). On the nonexistence of ancillary statistics. *Statist. Probab. Lett.* **15** 357–360. MR1193892 [https://doi.org/10.1016/0167-7152\(92\)90153-V](https://doi.org/10.1016/0167-7152(92)90153-V)
- [14] RIZZO, M. L. and SZÉKELY, G. J. (2010). DISCO analysis: A nonparametric extension of analysis of variance. *Ann. Appl. Stat.* **4** 1034–1055. MR2758432 <https://doi.org/10.1214/09-AOAS245>
- [15] RUZSA, I. Z. and SZÉKELY, G. J. (1982). Intersections of traces of random walks with fixed sets. *Ann. Probab.* **10** 132–136. MR0637381
- [16] RUZSA, I. Z. and SZÉKELY, G. J. (1983). Convolution quotients of nonnegative functions. *Monatsh. Math.* **95** 235–239. MR0712423 <https://doi.org/10.1007/BF01352002>

- [17] RUZSA, I. Z. and SZÉKELY, G. J. (1985). No distribution is prime. *Z. Wahrsch. Verw. Gebiete* **70** 263–269. [MR0799149 https://doi.org/10.1007/BF02451431](https://doi.org/10.1007/BF02451431)
- [18] RUZSA, I. Z. and SZÉKELY, G. J. (1988). *Algebraic Probability Theory*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Wiley, Chichester. [MR0974112](https://doi.org/10.1007/BF02451431)
- [19] SZÉKELY, G. J. (1986). *Paradoxes in Probability Theory and Mathematical Statistics. Mathematics and Its Applications (East European Series)* **15**. D. Reidel Publishing Co., Dordrecht. Translated from the Hungarian by Márta Alpár and Éva Unger. [MR0880020 https://doi.org/10.1007/978-94-017-2533-0](https://doi.org/10.1007/978-94-017-2533-0)
- [20] SZÉKELY, G. J. (2005). Half of a coin: Negative probabilities. Invited paper. *Wilmott Magazine* **66–68**.
- [21] SZÉKELY, G. J. (2006). Student's t-test for scale mixtures. In *The Second Erich L. Lehmann Symposium, May 19–22, 2004, Rice Univ.* (J. Rojo, ed.). *IMS Lecture Notes—Monograph Series* **49** 10–18.
- [22] SZÉKELY, G. J. and RIZZO, M. L. *The Energy of Data and Distance Correlation*. Chapman & Hall/CRC. Expected to be available in 2023.
- [23] SZÉKELY, G. J. and RIZZO, M. L. (2007). The uncertainty principle of game theory. *Amer. Math. Monthly* **114** 688–702. [MR2354439 https://doi.org/10.1080/00029890.2007.11920460](https://doi.org/10.1080/00029890.2007.11920460)
- [24] SZÉKELY, G. J. and RIZZO, M. L. (2009). Brownian distance covariance. *Ann. Appl. Stat.* **3** 1236–1265. [MR2752127 https://doi.org/10.1214/09-AOAS312](https://doi.org/10.1214/09-AOAS312)
- [25] SZÉKELY, G. J. and RIZZO, M. L. (2013). Energy statistics: A class of statistics based on distances. *J. Statist. Plann. Inference* **143** 1249–1272. [MR3055745 https://doi.org/10.1016/j.jspi.2013.03.018](https://doi.org/10.1016/j.jspi.2013.03.018)
- [26] SZÉKELY, G. J. and RIZZO, M. L. (2014). Partial distance correlation with methods for dissimilarities. *Ann. Statist.* **42** 2382–2412. [MR3269983 https://doi.org/10.1214/14-AOS1255](https://doi.org/10.1214/14-AOS1255)
- [27] SZÉKELY, G. J. and RIZZO, M. L. (2017). The energy of data, invited paper. *Annu. Rev. Stat. Appl.* **4** 447–479.
- [28] SZÉKELY, G. J., RIZZO, M. L. and BAKIROV, N. K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. [MR2382665 https://doi.org/10.1214/009053607000000505](https://doi.org/10.1214/009053607000000505)