Snapshots of modern mathematics
from Oberwolfach

№ 9/2026

Game-theoretic statistics

Glenn Shafer

Statistical modelling is the art of connecting abstract models with scientific or practical questions, so that the models' probabilities throw light on these questions. Standard statistical theory summarizes the results with p-values. Unfortunately, p-values are widely misunderstood. Game-theoretic statistics replaces them with e-values, also called betting scores. The model makes a forecast, the statistician makes an imaginary bet against it, and the e-value is the factor by which she multiplies her stake.

1 Testing by betting

If you put one euro on the table and use it to make a fair bet, what do you expect to get back? If the bet is all-or-nothing – as when the bet is settled by tossing a coin – you will get back two or zero euros. But what do you *expect*? The concept of "expected value" provides one way of summarizing your expectation. The *expected value* is found by multiplying each amount you might get back by its probability and adding up all these terms. Traditionally, mathematicians say that a bet is *fair* if the expected value is equal to the amount you put on the table. So the all-or-nothing bet is fair if the probabilities of two and zero are both 1/2, so that

$$\frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 0 = 1.$$

In this case we often say that the coin itself is fair.

Here is a more complicated example. Suppose you bet on how many tosses of a fair coin will be needed to get the first heads. Probability theory tells us that the probability that the first heads appears on the first toss is $1/2$, that the probability that it appears on the second toss is $(1/2)^2$, and so on. In general, the probability that the first heads appears on the nth toss is $(1/2)^n$. Suppose your bet requires you to pay one euro and pays back $9^{n-1}/5^n$ euros if the first heads appears on the nth toss. This bet is fair, because the formula $\sum_{n=1}^{\infty}(9/10)^{n-1} = 10$ implies that the expected value is

$$\sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n \cdot \frac{9^{n-1}}{5^n} = \sum_{n=1}^{\infty}\frac{9^{n-1}}{10^n} = \frac{1}{10}\cdot\sum_{n=1}^{\infty}\left(\frac{9}{10}\right)^{n-1} = \frac{1}{10}\cdot 10 = 1.$$

Even though it is fair, the bet might have a very surprising outcome. Suppose the first heads appears on the 100th toss. Then the payback from your investment of one euro is $9^{99}/5^{100}$. This is more, in euros, than a 3 followed by 24 zeros. This shows that the bet is entirely imaginary, because the wealth of the entire world is minuscule in comparison. It also makes us question the hypothesis that assigns probability $(1/2)^n$ to the first heads appearing on the nth toss. *Perhaps this hypothesis does not provide a good forecast of what will happen if the experiment is repeated.*

Why would a statistician ever think about this imaginary bet and its payoff $9^{n-1}/5^n$? She would have good reason to think about the number $9^{n-1}/5^n$ if she thought that the probability of heads on each toss is really $1/10$ instead of $1/2$. Under this *alternative hypothesis*, the probability for seeing $n-1$ tails followed by a heads is $(9/10)^{n-1}(1/10)$ instead of $(1/2)^n$, the probability assigned to this event by the original hypothesis. The ratio of these two probabilities,

$$\frac{(9/10)^{n-1}\cdot(1/10)}{(1/2)^n} = \frac{9^{n-1}}{10^n}\cdot 2^n = \frac{9^{n-1}}{5^n},$$

is called the *likelihood ratio* for comparing the two hypotheses. Many statisticians consider the likelihood ratio by itself a good measure of the success of the alternative hypothesis relative to the original hypothesis [2, 12]. Our picture supports the meaningfulness of the likelihood ratio by recognizing it as the payoff of a bet against the original hypothesis.

Waiting so long for the first heads might make our statistician (let's call her *Statistician* with a capital S) question even her alternative hypothesis, because it assigns a probability less than 0.000033 to the first 99 tosses being tails. But a less extreme outcome might confirm Statistician's alternative hypothesis. For example, if the first heads appears on the 10th toss, then the payoff is $9^9/5^{10} \approx 40$.

Multiplying your capital by 40 is not like multiplying it by 3 followed by 24 zeros. But it is already very strong as evidence against the original hypothesis

that the probability of heads is $1/2$. In fact, it is what you might expect under the alternative hypothesis that the probability of heads is only $1/10$. If you have studied probability theory, you might know that under Statistician's alternative hypothesis that the probability of heads is $1/10$, the number n of tosses to get the first heads has a geometric distribution with expected value 10. This implies that the logarithm of the payoff,

$$\ln\left(\frac{9^{n-1}}{5^n}\right) = (n-1)\ln 9 - n\ln 5,$$

has expected value $9\ln 9 - 10\ln 5$, which is exactly the logarithm of the payoff when $n = 10$. In Section 5, we will discuss why it is natural to make the comparison on the logarithmic scale.

2 Test martingales

Instead of imagining a single bet at the outset, Statistician might imagine a sequence of bets, one on each toss.

1. First she pays 1, getting back $9/5$ if the first toss comes out tails and $1/5$ if it comes out heads.
2. She stops betting if the toss came out heads. If it came out tails, she pays the $9/5$ she now has to get $(9/5)^2$ if the first toss comes out tails and $(9/5)\cdot(1/5)$ if it comes out heads.
3. And so on. She stops the first time she gets a head. But if she has all tails after k tosses, she bets her capital at that point, $(9/5)^k$, to get back $(9/5)^{k+1}$ if the $(k+1)$st toss comes out tails and $(9/5)^k \cdot (1/5)$ if it comes out heads.

The net result will be the same as the first bet we studied: Statistician begins with 1 and ends up with $(9/5)^{n-1} \cdot (1/5)$, where n is the number of tosses to get a heads.

This type of random sequence, where the results of successive bets are compounded, is called a *test martingale*.[1] It turns out that any bet on a sequence of outcomes that uses unit capital and necessarily results in a nonnegative payoff can be obtained by a test martingale.

3 Two ways of testing: e-values and p-values

A nonnegative payoff that has expected value one or less under the hypothesis being tested is called an *e-value* or *betting score*. As we have just seen, a test

[1] For centuries, betting strategies in casinos were called martingales. Now mathematicians use the name for random sequences of capital that result from betting strategies [9].

martingale is the result of multiplying successive e-values.[2]

In order to use e-values and test martingales for statistical inference, Statistician must adopt the *principle* that a large e-value, for a bet chosen in advance, discredits the hypothesis. It should be kept in mind that this is a principle, not a theorem. How large is large? This is a matter of judgement. Some statisticians have established conventions about how to describe the degree of discredit associated with different e-values. One convention is that 4 is serious evidence, 10 is strong evidence, and multiples of 10 are very convincing. The value of such a convention depends on the context, however. A large e-value can be taken seriously only if it is based on plausible alternatives that Statistician has good reason to consider. Otherwise, it can easily be dismissed as a lucky stab in the dark or even as a likely mistake.[3]

The evidentiary role of an e-value is analogous to that of the better known statistical concept of a "p-value". To define a p-value, you first choose a test statistic, or function of the data; the *p-value* for testing a hypothesis is the hypothesis's probability for the test statistic being as large or larger than it actually turned out to be. When the p-value comes out very small, the hypothesis is discredited. We used a p-value in Section 1. There, the test statistic was the number of tosses required to get the first heads, and the hypothesis was Statistician's alternative hypothesis, which gave probability $9/10$ to heads on each toss. We supposed that the first heads came on the 100th toss. So the p-value was the hypothesis's probability for the first heads coming only on the 100th toss or later, and this was a very small number — only about 0.000033.

Statisticians have used p-values for two centuries, and they remain more widely used than e-values by a huge margin. But e-values have several advantages. As we have seen, we can combine evidence from successive observations by multiplying e-values. The average of two or more e-values, possibly weighted, is also an e-value, and this provides a way of combining evidence based on different alternatives to the hypothesis being tested. If one of the bets produces a sufficiently large e-value, the average e-value will also be large. Combining successive observations or distinct tests is possible for p-values but more awkward. One way to combine p-values is to convert them into e-values. There are many

[2] Jean Ville studied nonnegative martingales starting with unit capital in the 1930s [15]. Herbert Robbins and his collaborators used composite nonnegative supermartingales in statistics beginning in the 1960s [1]. Vladimir Vovk emphasized the direct interpretation of high values of a test martingale in the early 1990s [16]. The simpler concept of an e-value emerged in the late 2010s [4, 5, 17, 19]. The name *e-value* is due to Vovk and Ruodu Wang [17]. Unfortunately, there are several other established uses for "e-value" in various specialized fields of statistics. This provides a reason for using the alternative name *betting score* [13].

[3] In addition to testing, there are other applications of e-values. In some applications, modest e-values are useful [6]. In some others, only very large e-values are considered [18].

functions that do this; the simplest may be $\sqrt{1/p} - 1$. It converts the p-value 0.05, conventionally considered serious evidence, to the e-value 3.47.

In addition to testing a single hypothesis about the probabilities for an event, statisticians sometimes simultaneously test multiple hypotheses about the probabilities for the event. In this case, they say that they are testing a *composite hypothesis*. An e-value for testing a composite hypothesis is a payoff that has expected value one or less under each of the hypotheses. A p-value is the maximum or supremum of the different hypotheses' p-values using the same test statistic.

4 Free play and a fundamental principle

If Statistician tested the fairness of her coin using the martingale we described in Section 2, and the first heads appeared on the 10th toss, then she saw this sequence of cumulative e-values:

$$\frac{9}{5},\ \left(\frac{9}{5}\right)^2,\ \left(\frac{9}{5}\right)^3,\ \left(\frac{9}{5}\right)^4,\ \left(\frac{9}{5}\right)^5,\ \left(\frac{9}{5}\right)^6,\ \left(\frac{9}{5}\right)^7,\ \left(\frac{9}{5}\right)^8,\ \left(\frac{9}{5}\right)^9,\ \left(\frac{9}{5}\right)^9 \cdot \left(\frac{1}{5}\right)$$

After the fourth toss, she already had an e-value of $(9/5)^4$, which is greater than 10. At that point, she might have felt that this was all the evidence against the coin's fairness that she needed. So she might have stopped. Is this legitimate? Yes. But Statistician's right to stop and count $(9/5)^4$ as evidence is again a *principle*, not a theorem. It has been called the *principle of optional stopping*.

Here is a more general principle that is equally reasonable.

> **Fundamental principle for testing by betting.** If you make successive bets against a forecaster, beginning with a unit stake and never risking additional capital, and your capital becomes large, then you discredit the forecaster.

The condition "never risk additional capital" means that on each round of betting, Statistician can risk only her current cumulative capital: her initial unit capital plus gains so far minus losses so far. She cannot make a bet that might make her cumulative capital negative.

This fundamental principle not only allows Statistician to stop when she wants, it also allows her to begin testing with no plan for when to stop or how to bet on later rounds if she does not stop. It even allows her to change the experiment that produces the outcomes she is betting on. In short, it allows *free continuation*.

5 Statistician's logarithmic utility

The difference between the two e-values 5 and 20 seems very important, while the difference between 105 and 120 seems negligible. The increase from 5 to 20 is a quadrupling. For a comparable impact starting from 105, perhaps we should quadruple it to 420.

When the difference between perceptions depends on ratios instead of differences in this way, we say that the perception is logarithmic.[4] This is because multiplication on the original scale is addition on the logarithmic scale. In general, $\ln(xy) = \ln(x) + \ln(y)$. So $\ln(4 \cdot 5) = \ln 4 + \ln 5$ and $\ln(4 \cdot 105) = \ln 4 + \ln 105$; in both cases, $\ln 4$ is being added.

A function that tells how much a person values different amounts of money is called her *utility function*. One way to choose among actions whose payoffs depend on an outcome is to maximize your expected value for your utility function. Assuming that Statistician's utility function is the logarithm, we can recommend that she bet so as to maximize her expected value for the logarithm of her e-value. Because this was advocated by John L. Kelly, Jr., in 1956 [7], it is called the *Kelly criterion*.[5]

A bit of notation may help here. Statistician is betting against a probability distribution P for some outcome Y. She has her own probability distribution Q for Y. She bets by paying 1 for a payoff S that is a function of Y and has expected value 1: $\mathbf{E}_P(S(Y)) = 1$. Maximizing her expected utility means choosing S to maximize $\mathbf{E}_Q(\ln(S(Y)))$. As it turns out, this maximum is achieved when S is the likelihood ratio: $S(Y) = {}^{Q(Y)}/_{P(Y)}$. See [13] for a proof.[6] As we mentioned in Section 1, many statisticians already advocate the likelihood ratio as a measure of evidence.

In the example in Section 1, the outcome Y on each toss of the coin is either heads or tails. The hypothesis P assigns these possible outcomes probabilities $P(\text{heads}) = P(\text{tails}) = {}^1/_2$. Statistician's alternative Q assigns them probabilities $Q(\text{heads}) = {}^1/_{10}$ and $Q(\text{tails}) = {}^9/_{10}$. So on each toss Statistician's Kelly bet produces the e-value

$$\frac{Q(\text{heads})}{P(\text{heads})} = \frac{({}^1/_{10})}{({}^1/_2)} = {}^1/_5$$

[4] Human perception of many stimuli is approximately logarithmic. The phenomenon is called *Fechner's law*. It has been observed in the perception of weight, noise, light, smell, etc.

[5] The Kelly criterion has also been used in real gambling and financial risk taking. The criterion requires that the gambler or investor has a probability distribution, but it does not require their betting opportunities to be based on a forecaster's probabilities [3, Ch. 10].

[6] For extensions to composite hypotheses, see [10, 11].

or the e-value

$$\frac{Q(\text{tails})}{P(\text{tails})} = \frac{(9/10)}{(1/2)} = 9/5.$$

When she multiplies these e-values for successive tosses, she obtains the test martingale we studied in Section 2. The quantity $\mathbf{E}_Q(\ln(S(Y)))$ tells her how well she can expect to do on each toss:

$$\mathbf{E}_Q(\ln(S(Y))) = \frac{1}{10}\ln\left(\frac{1}{5}\right) + \frac{9}{10}\ln\left(\frac{9}{5}\right) \approx 0.368.$$

These expected values add. So Statistician's expected log betting score for 10 tosses is about 3.68, corresponding to a betting score of about $e^{3.68} \approx 39.7$. This is close to the betting score of 40 that we found at the end of Section 2 for the case where the first heads comes on the 10th toss.

Why not choose S to maximize $\mathbf{E}_Q(S(Y))$ instead of $\mathbf{E}_Q(\ln(S(Y)))$? For one thing, it is too risky. In our example, the bet S that maximizes $\mathbf{E}_Q(S(Y))$ is

$$S(\text{heads}) = 0 \text{ and } S(\text{tails}) = 2.$$

Statistician's probability distribution Q tells her that if she adopts this bet, she has probability $1/10$ of losing all her money and not being able to bet on a second toss. In examples, where Q gives small probabilities to all possible outcomes, maximizing $\mathbf{E}_Q(S(Y))$ is even more dangerous. It requires that Statistician stakes all her capital on the outcome with the greatest of these small probabilities. The possible payoff will be large but the probability of losing everything will also be very large.

6 Statistical modelling

As we have seen, e-values can be used to test forecasters who give probabilities for outcomes that we will observe. They can also be used to make inferences about facts that we will never observe and questions that we do not expect to settle with absolute certainty and precision. What is the mass of Jupiter? Does a particular medical treatment do any good? To address such a question with e-values, Statistician must make an argument for connecting a probability distribution with the question. Statistical modelling is the art of making such arguments.

A classical example of statistical modelling is the theory of errors developed by Laplace and Gauss in the early 19th century. Suppose Statistician thinks that a normal probability distribution with mean zero and variance one (the familiar "bell-shaped curve" of statistics) has described past errors of her measuring instrument reasonably well. She argues that this is reason enough to consider it

a good forecast of the error the instrument will make measuring a quantity μ. She will observe only the measurement, not the error, because she does not know μ. But she imagines a game in which one player (let's call her *Oracle*) does know μ and announces it at the outset to two other players, *Forecaster* and *Skeptic*. The game has N rounds of play, and every round n ends with the players observing a measurement y_n of μ. Just before observing y_n, Forecaster announces a probability distribution for y_n, and Skeptic bets against it. Here Statistician does not allow Forecaster and Skeptic to be free players. Instead, she assigns them strategies.

1. She tells Forecaster to announce the normal distribution with mean μ and variance 1 on every round. Call it P.
2. She assigns Skeptic a strategy that uses his knowledge of $y_1, \ldots, y_{n-1}$ and μ to bet against P on every round n. The strategy might, for example, use the Kelly bet Q/P, where Q is with the normal distribution with mean $(y_1 + \cdots + y_{n-1})/(n-1)$ and variance 1. In this case, the alternative Q does not encode Skeptic's opinion. It simply uses the previous observations $y_1, \ldots, y_{n-1}$ to challenge P as a forecast of y_n.

Because Statistician does not know μ, she also does not know what moves by Skeptic and Forecaster result from the strategies she has assigned them. But she does observe the ys. So she can calculate Skeptic's final capital as a function of μ. Because she thinks the normal distribution she has assigned to Forecaster is a good forecast, she does not think the final capital will be large. This does not allow her to identify the value of μ precisely, but it may allow her to rule out values for which the final capital is too large. If she sets a threshold and rules out the values of μ for which the final capital would exceed that threshold, the remaining values of μ constitute what statisticians call a *confidence set* for μ. The game-theoretic justification of this confidence set lies purely in Statistician's strong but fallible confidence that Skeptic will not defeat Forecaster's forecasts.

Pseudo-random numbers generated by computers provide another source of well-tested probability forecasts that can be used in statistical modelling [8]. One of the simplest and best known examples of their use is survey sampling. In survey sampling, Statistician finds a way to label the members of a population with numbers, so that those whose numbers are randomly selected by the computer can be queried for some feature. For example, if the population consists of people, the feature might be presence of a particular gene or possession of a high school diploma. When an individual is sampled from the population, it is recorded as 1 if it has the feature, 0 if not. The resulting sequence of zeros and ones, say $y_1, \ldots, y_N$, constitutes the observations. To estimate the unknown proportion θ of the population with the feature, Statistician imagines that Oracle first announces θ. Then, before each successive y_n is observed, Forecaster uses θ as his forecast for y_n. By doing so, he is announcing a

probability distribution P: $y_n = 1$ is assigned probability θ and $y_n = 0$ is assigned probability $(1 - \theta)$. Then Skeptic makes a bet against P using the previous observations, as in the measurement example. Again, Statistician can form a confidence set for θ based on the assumption that Forecaster is making good forecasts.

In these two examples, measurement and random sampling, game-theoretic statistics gives us a new way of understanding established statistical methods. In more complicated problems, the game-theoretic perspective has also led to new methods; some are described in [14, Ch. 10] and [11].

7 Game-theoretic probability

As a branch of mathematics, game theory studies strategies in fully defined games — games for which the mathematician specifies the players' goals and the extent to which they see other players' moves. In real games, players might also have private information and might acquire more private information as play proceeds, but in the mathematical theory a player's strategy can use only the previous moves in the game that the player sees.

We leave the realm of pure mathematics when we allow Forecaster and Skeptic (or Statistician, who is pulling their strings) to play freely instead of following a strategy specified at the outset of play. We also sometimes find it useful to introduce other players who can play freely, including Oracle and also Experimenter, who may change the experiment for the next round of play or announce auxiliary information, as in regression or classification problems.

It is also interesting and useful, however, to re-enter the realm of mathematics by studying strategies for Forecaster and Skeptic. This is *game-theoretic probability*.[7] To fully define the game, we need a player who announces outcomes; we call this player Reality. We can then ask whether Skeptic has a winning strategy for a particular goal.

Consider, for example, this game between Forecaster, Skeptic, and Reality, which continues for 1,000 rounds. Skeptic begins with one euro. Forecaster begins each round by announcing a probability for heads, then Skeptic bets against the forecast, and then Reality says "heads" or "tails". Skeptic must bet in such a way that his current capital (his initial unit capital plus any winnings minus any losses) never becomes negative, no matter what Reality does. Write $\overline{p}$ for the average of Forecaster's 1,000 probabilities, $\mathcal{K}$ for Skeptic's capital at the end of play, and #heads for number of times Reality says "heads". And set

[7] The most thorough study of game-theoretic probability is [14].

Skeptic this goal:

$$\text{Either } \left| \frac{\#\text{heads}}{1{,}000} - \overline{p} \right| < 0.1 \text{ or else } \mathcal{K} \geq 10.$$

Skeptic has a winning strategy in this game. In other words, Skeptic can achieve his goal no matter what Forecaster and Reality do. When Skeptic follows his winning strategy, either the frequency of heads will come within ten percentage points of Forecaster's average p, or else Skeptic's final capital will be at least 10.

The theorem stating that Skeptic has a winning strategy in this game generalizes Chebyshev's version of the law of large numbers to the case where Forecaster is a free player. This is only one game-theoretic instance of the many-faceted *law of large numbers*, but it illustrates how discrete-time probability theory can be translated into game theory and then generalized.

The game-theoretic law of large numbers also supports and generalizes an argument made by Kelly in 1956 for Skeptic using logarithmic utility. Laws of large numbers, game-theoretic or not, are available only when we are adding. Logarithms and their expected values add, and in game theoretic probability this feature extends to the case where Forecaster is a free player. This does not work if Skeptic tries to maximize the expected value of some other function of his betting score on each round – the logarithm is crucial.

8 How to discourage real gambling

Probability theory began as a theory of gambling in games of pure chance. When Jacob Bernoulli, who proved the first law of large numbers, undertook to apply it to civil, criminal, and business affairs, he minimized the connection with gambling. Most statisticians have followed his example. We want statistical inference to live in the realm of reason, not in the casino.

The mathematical statisticians who have been developing game-theoretic statistics continue to work in reason's realm. They do not seek to promote gambling. Some even advocate its legal prohibition. But the value of game-theoretic statistics lies in the insights that can be gained by making the betting game between Forecaster and Skeptic explicit, and this makes hiding the connection with gambling impossible. Instead of trying to hide it, we should use it to educate the public about the perils of gambling.

There is little peril in playing *nonnegative* martingales. Even if the euro you are risking is real rather than imaginary, you cannot bankrupt yourself by trying to multiply it without risking more. But people can destroy their lives when, consciously or unconsciously, they play martingales that can become negative, and deeply negative. It has become easier and easier to play such martingales, whether by impulsively making another online sports bet or by

speculating in financial securities on margin. It should be the duty of teachers of probability and statistics to teach how easily gamblers and day traders can delude themselves.

Acknowledgments

Game-theoretic statistics is a very new field. Its excitement and promise are matched by the diversity of viewpoints within the community of researchers developing it. This snapshot, which emphasizes betting rather than the mathematics of e-values and related concepts, owes the most to my long-time collaborator Vladimir Vovk. I am also grateful for insights provided by Oberwolfach's Workshop 2419b, held in May 2024, and especially by its organizers, Peter Grünwald, Aaditya Ramdas, Ruodu Wang, and Johanna Ziegel.

References

[1] D. A. Darling and H. Robbins, *Confidence sequences for mean, variance, and median*, Proceedings of the National Academy of Sciences **58** (1967), no. 1, 66–68.

[2] A. W. F. Edwards, *Likelihood*, Cambridge, 1972.

[3] S. N. Ethier, *The doctrine of chances: Probabilistic aspects of gambling*, Springer, 2010.

[4] P. Grünwald, R. de Heide, and W. M. Koolen, *Safe testing (with discussion)*, Journal of the Royal Statistical Society, Series B **86** (2024), no. 5, 1091–1171.

[5] S. R. Howard, A. Ramdas, J. McAuliffe, and J. Sekhon, *Time-uniform Chernoff bounds via nonnegative supermartingales*, Probability surveys **17** (2020), 257–317.

[6] N. Ignatiadis, R. Wang, and A. Ramdas, *E-values as unnormalized weights in multiple testing*, Biometrika **111** (2023), no. 2, 417–439.

[7] J. L. Kelly Jr., *A new interpretation of information rate*, The Bell System Technical Journal **35** (1956), no. 4, 917–926.

[8] R. T. Kneusel, *Random numbers and computers*, Springer, 2018.

[9] L. Mazliak and G. Shafer (eds.), *The splendors and miseries of martingales: Their history from the casino to mathematics*, Birkhäuser, 2022.

[10] A. Ramdas, P. Grünwald, V. Vovk, and G. Shafer, *Game-theoretic statistics and safe anytime-valid inference*, Statistical Science **38** (2023), no. 4, 576–601.

[11] A. Ramdas and R. Wang, *Hypothesis testing with e-values*, Foundations and Trends in Statistics **1** (2025), no. 1-2, 1–390.

[12] R. Royall, *Statistical evidence: A likelihood paradigm*, Chapman & Hall, 1997.

[13] G. Shafer, *Testing by betting: A strategy for statistical and scientific communication (with discussion)*, Journal of the Royal Statistical Society, Series A **184** (2021), no. 2, 407–478.

[14] G. Shafer and V. Vovk, *Game-theoretic foundations for probability and finance*, Wiley, 2019.

[15] J. Ville, *Étude critique de la notion de collectif*, Gauthier-Villars, 1939.

[16] V. Vovk, *A logic of probability, with applications to the foundations of statistics (with discussion)*, Journal of the Royal Statistical Society, Series B **55** (1993), no. 2, 317–351.

[17] V. Vovk and R. Wang, *E-values: Calibration, combination and applications*, The Annals of Statistics **49** (2021), no. 3, 1736–1754.

[18] R. Wang and A. Ramdas, *False discovery rate control with e-values*, Journal of the Royal Statistical Society, Series B **84** (2022), no. 3, 822–852.

[19] L. Wasserman, A. Ramdas, and S. Balakrishnan, *Universal inference*, Proceedings of the National Academy of Sciences **117** (2020), no. 29, 16880–16890.

Glenn Shafer *is university professor emeritus at Rutgers University.*

License
Creative Commons BY-SA 4.0

Mathematical subjects
Probability Theory and Statistics

DOI
10.14760/SNAP-2026-009-EN

Connections to other fields
Computer Science, Engineering and Technology, Finance, Humanities and Social Sciences, Life Science

---

Snapshots of modern mathematics from Oberwolfach provide exciting insights into current mathematical research. They are written by participants in the scientific program of the Mathematisches Forschungsinstitut Oberwolfach (MFO). The snapshot project is designed to promote the understanding and appreciation of modern mathematics and mathematical research in the interested public worldwide. All snapshots are published in cooperation with the IMAGINARY platform and can be found on www.imaginary.org/snapshots and on www.mfo.de/snapshots.

ISSN 2626-1995

Junior Editor
Noam Szyfer
junior-editors@mfo.de

Senior Editor
Anja Randecker
senior-editor@mfo.de

Mathematisches Forschungsinstitut
Oberwolfach gGmbH
Schwarzwaldstr. 9–11
77709 Oberwolfach
Germany

Director
Gerhard Huisken

Mathematisches
Forschungsinstitut
Oberwolfach

IMAGINARY
open mathematics