

On Testing by Imaginary Betting

by *Glenn Shafer*

Data analysis requires principles as well as mathematics. Traditionally, we have relied on Cournot's principle when we use probability theory for data analysis. But when we test by betting instead of relying on small probabilities, we can formulate principles that dig deeper into statistical practice and apply more broadly.

“An event with very small probability will not happen.” “An event with probability close to one will happen.” These aphorisms have gone by various names, including *Cournot's principle*, the *principle of improbability*, and the *principle of practical certainty*.

Cournot's principle does a lot of work. We predict events that have very high probability. We use the happening of events with very small probabilities to discredit theories that attribute the small probabilities to them. (Here the small probabilities are called *p-values*.) Augustin Cournot, depicted in Figure 1, told us that it is the *only* way probability theory makes contact with the world of phenomena.

But something is wrong. Events of small probability happen all the time. One of the tickets in a lottery will win. When a probability distribution is continuous, all the possible outcomes have zero probability. How do we discipline Cournot's principle? Are there further principles for its use?

Testing by betting provides a framework for formulating such principles. Instead of testing probabilities, we test a forecaster. The forecaster might be a theory, a person, or an artificial intelligence. We interpret his forecasts as betting offers. Perhaps he states probabilities and offers to sell random variables for their expected values. Perhaps he makes more limited betting offers. You bet against him by deciding which offers to accept. But the bets are imaginary. The factor by which you multiply the capital you risk is your *betting score* or *e-value*. You can make successive bets, using the capital from one bet to make the next. Your cumulative capital after multiple rounds is your *cumulative betting score*.

In this framework, we can refine and expand Cournot's principle into five fundamental principles.

- 1. Principle for testing by betting.** If you make successive bets against a forecaster, beginning with a unit stake and never risking additional capital, and you obtain a large cumulative betting score, then you discredit the forecaster.

A large betting score has a small probability of happening. Our principle for testing by betting generalizes the use of p-values to discredit probabilities, because the forecaster might not state a joint probability distribution for the successive outcomes. But it also disciplines the practice, because you must bet before you see the outcomes.

How large does a cumulative betting score need to be in order to discredit a forecaster? It depends. The degree of discredit depends in part on the reputation of the players. Cautious betting based on expertise and extending over many rounds of betting is more persuasive than careless betting based merely on magical thinking.

2. **Principle for multiple discredit.** If you simultaneously test multiple forecasters who are forecasting the same outcome, obtaining a large betting score against each forecaster you test, then you have discredited all the forecasters.

Why? Because if you multiply each unit stake by a large factor, you have multiplied your total stake by a large factor. This principle reduces the problem of testing statistical models to the problem of testing a single forecaster.

3. **Cournot's principle, disciplined.** You may predict that a forecaster who has consistently withstood certain test strategies in the past will similarly withstand a similar test strategy in the future.

A strategy for successive betting determines how your evolving capital depends on forecasts and outcomes. This evolving capital as a function of forecasts and outcomes is called a *supermartingale*. If the strategy does not risk the capital becoming negative, the strategy is called a *test strategy*, and the supermartingale is nonnegative. When a nonnegative supermartingale has a large value unless an event happens, the event has small (*upper*) *probability*.

Our version of Cournot's principle reduces the aphorism that events of small probability will not happen to a type of philosophical induction. The future will be like the past. The principle is disciplined, because it applies only to test strategies that have been previously used against the forecaster.

4. **Principle for point prediction.** If you can replicate a payoff with a betting strategy using bets that the forecaster has withstood in the past, you may use the cost of the replication as a prediction of the payoff.

This principle generalizes and disciplines the practice of using expected values as predictions. We call predictions based on Principles 3 or 4 *warranted*.

5. **Principle for statistical estimation.** A warranted prediction about outcomes that are not observed may further warrant the inference that related unknowns are consistent with the prediction.

Suppose, for example, that you make a prediction about the errors of a measuring instrument based on experience testing the instrument to measure known quantities. If your prediction is about the errors in measuring an unknown quantity, then you will not observe the errors you are predicting, but your prediction of the average error to be within certain bounds will warrant the inference that the unknown quantity you are measuring is within certain bounds.

Probability theory was originally a theory about betting, and it is still driven by intuitions about betting. Trying to appear more objective, statistical theory has often tried to hide this underlying logic of betting. The five principles listed here show that by bringing betting back to the surface, we are able to articulate aspects of data analysis that are often hidden from students of statistics and from the public.

References

- [1] L. Mazliak and G. Shafer, “The Splendors and Miseries of Martingales: Their History from the Casino to Mathematics”, Cham, Switzerland: Birkhäuser, 2022.
- [2] G. Shafer, “Testing by betting: A strategy for statistical and scientific communication”, Journal of the Royal Statistical Society: Series A, vol. 184, no. 2, pp. 407–478, 2021.

Please contact:

Glenn Shafer, Professor Emeritus, Rutgers University, USA
gshafer@business.rutgers.edu



Figure 1: Antoine Augustin Cournot (1801-1877). He is celebrated by economists for his work on duopoly, but his contribution to statistics is underrated. As the philosopher Fernand Faure wrote in 1905, it was too philosophical for the statisticians, too statistical for the philosophers.