Many seemingly robust findings in particle physics and astronomy have crumbled while less convincing data have brought glory. CERN is hoping for the latter in its quest for the Higgs boson

CERN's Gamble Shows Perils, Rewards of Playing the Odds

"It would be unprecedented in the history of science," says Michael Riordan, a particle physicist at the Stanford Linear Accelerator Center (SLAC). "It would be the greatest signal since the discovery of quarks, and they can't chase it down."

Riordan is referring to a last-gasp chase by physicists at CERN, the European particle physics laboratory in Switzerland, for solid evidence of a particle known as the Higgs boson. Experiments using CERN's Large Electron-Positron collider (LEP) turned up hints of the elusive Higgs particle just before LEP's scheduled shutdown (*Science*, 22 September, p. 2014). The tantalizing results won LEP a monthlong stay of execution, but nobody expects CERN to nail down whether the discovery is real or a mirage. "Oh, no—that is quite excluded," says Peter Igo-Kimenes, a particle physicist in charge of combining data from the four LEP experiments.

Igo-Kimenes is certain that even an extra month of experiments—about as much overtime as LEP can get without triggering harsh penalty clauses in builders' contracts—will not boost the data across the threshold particle physicists use to separate true discoveries from the chaff of statistical fluctuations: five standard deviations, or five sigma. The LEP data are languishing in the three- to four-sigma range, far short of what is needed to declare a stone-solid discovery.

The endgame drama at CERN has focused attention on just what it takes to stake a claim to a particle or an event in areas of science where the data are fuzzy, sightings are fleeting, and probabilities rule. Glimpse your new fundamental particle or extrasolar planet at the right moment, with the right degree of confidence, and you win the discoverer's laurels. Otherwise, you are just another precursor or confirmer, fodder for footnotes. Or, if you are really unlucky, your seemingly robust result will turn out to be a product of experimental bias, and you may wind up humiliated in front of your peers.

To physicists and astronomers, the five-sigma rule is the acid test for judging discoveries and assigning credit. So why do the physicists at LEP persist when they know they can't possibly make the grade? Because they also know that reality is a lot messier than theory. In practice, the five-sigma rule is far from golden. Discoveries that seem statistically unassailable can vanish overnight, while flimsier looking findings have entered the award rosters and the textbooks without cavil. Qualitative factors, such as the reputation of a team of scientists, whether a finding conforms to prevailing theory, and how and why the team announces a discovery, can determine whether it wins the Nobel Prize or languishes as an also-ran.

Inspiration, perspiration, or ...? When "seeing" is statistical, credit for a discovery can seem a matter of divine whim.

ILLUSTRATION: T. SMITH

Vanishing probabilities

To a statistician, such vagaries may seem absurd. On the surface, finding a new particle should be little different from figuring out whether a medication is effective or when a coin is biased. Numerically, a five-sigma result corresponds to less than one chance in 3 million that a sighting is due to chance (see sidebar). Even a much weaker three-sigma result in particle physics means that the scientists are 99.9% sure that their signal didn't appear by accident. By definition, then, a mere one three-sigma result in 1000 results is wrong. Right?

Not exactly. "Half of all three-sigma results are wrong," says John Bahcall, a particle physicist and astrophysicist at Princeton University. "Look at the history." He's right: Not only do a surprising number of three-sigma results vanish on closer inspection, but an astounding number of five- and six-sigma results have done so, too.

In the mid-1980s, for example, physicists at the Organization for Heavy Ion Research (GSI) in Darmstadt, Germany, looked well on their way to the Nobel Prize. Two separate experiments had found peaks in their data, hinting at a new particle in the 600- to 700-KeV (thousand electron volts) range. It wasn't predicted by the Standard Model, but the signal was strong—more than six sigma, corresponding to a one-in-a-billion chance of error. Today, the mysterious particle is gone forever. "We have given up the experiments," sighs GSI physicist Helmut Bokemeyer. "We have not been able to see what we had seen before."

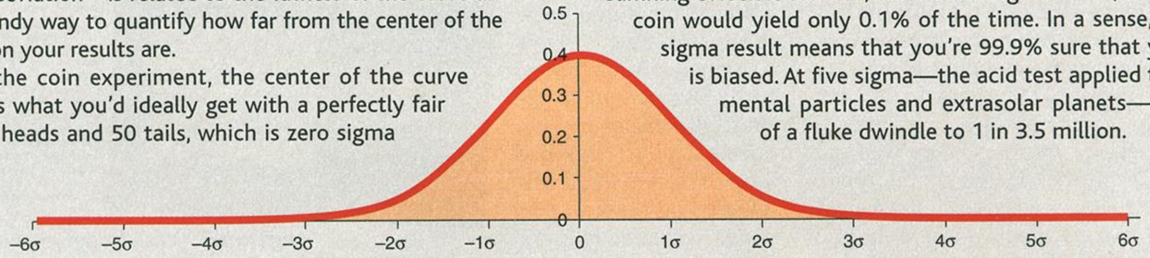
There was probably nothing to see in the first place. Experiments at Brookhaven, Argonne, and elsewhere tried finding similar peaks and found nothing, and the GSI result died in a firestorm of controversy (*Science*, 10 January 1997, p. 148). "That's a fairly sad episode," says Brookhaven's William Zajc.

What went wrong? Bokemeyer thinks that the GSI experiments were highly "opti-

A Greek Letter, Demystified

In principle, screening out statistical noise from a particle physics experiment is a lot like determining whether a coin is biased toward heads or tails. Random events such as coin flips and particle detections tend to follow a bell curve distribution. Sigma—the standard deviation—is related to the fatness of the curve and gives a handy way to quantify how far from the center of the distribution your results are.

With the coin experiment, the center of the curve represents what you'd ideally get with a perfectly fair coin—50 heads and 50 tails, which is zero sigma away from the center. As you rack up a larger and larger surplus of heads (or tails), you move away from the center of the curve. A result of 55 heads and 45 tails is one sigma away from the center; there's about a 16% chance that an unbiased coin will give that result. On the other hand, 60 heads and 40 tails is a two-sigma result, which has only about a 2% chance of happening with an unbiased coin. A run of 65 heads and 35 tails would be pretty damning evidence for bias; it's a three-sigma result, which a fair coin would yield only 0.1% of the time. In a sense, a three-sigma result means that you're 99.9% sure that your coin is biased. At five sigma—the acid test applied to fundamental particles and extrasolar planets—the odds of a fluke dwindle to 1 in 3.5 million. **–C.S.**



mized" to find the peak. In other words, change the experiment ever so slightly, and the peak disappears, which explains why the result is so difficult to reproduce. "Otherwise, we have no idea what it could be," he says. There are other possible explanations. For instance, the researchers might have begun an experimental run and looked for a growing peak to make sure that the equipment was set up properly. If there was no indication of a bulge in the data, they would change aspects of the experiment and try again.

Some physicists believe that this habitual restarting of the experiment may have introduced an unintentional bias into the results. Subtle statistical effects like this, or problems with equipment, or a slight error in calculation, or an overlooked source of conflicting data, can throw off statistical calculations in a tremendous way. "It's the systematic errors that kill you," Bahcall says. Bahcall knows that the perils of failure against the odds stretch far beyond particle physics: Seven years ago he saw it strike on a cosmic scale.

Vanishing planets

"It was the thing that one fears more than anything else in one's scientific life, and it was happening," says Andrew Lyne, an astrophysicist at Jodrell Bank Observatory in Manchester, U.K. "I certainly at the time thought that it was the end of my career."

In January 1992, Lyne was celebrating a monumental discovery. He and his team had spotted what appeared to be the first planet circling a foreign star. Their radio telescope had found a pulsar whose clocklike pulses sped up and slowed down in a way that suggested it was being tugged around by an invisible orbiting body. "Indeed, based upon a straightforward statistical analysis, the effect was very highly significant—hundreds of sigmas, a certainty," Lyne says. "We did all sorts of tests on the data and tried to think of all the possible ways we might be making a mistake." After finding their procedures sound, the team published their discovery: the first extrasolar planet. "It received a lot of interest, as you can imagine, from the media and others," he says.

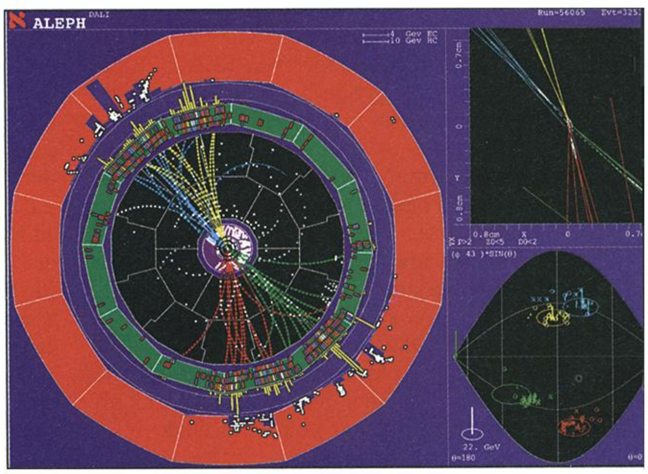
Bahcall, then president of the American Astronomical Society, called a special session together to discuss the discovery at the society's annual meeting. But then disaster struck. "Ten or 12 days before I was due to give that talk, I discovered the error and the true source for the periodicity," Lyne says. "It was rather subtle."

When timing signals that come from pulsars, astronomers have to correct for Earth's motion around the sun, which introduces a tiny periodic distortion in the signal. To save computer resources, Lyne's group used an approximation of Earth's orbit for the preliminary calculations. For a more detailed analysis, they planned to switch to a more accurate model and redo their work from scratch. Unfortunately, with one of the 200 or so pulsars that they looked at, they forgot to perform the more accurate calculation and based their conclusions on the rough approximation. "The full high-precision analysis was not carried out," Lyne says. The slight inaccuracy in accounting for Earth's orbit led to a periodic signal that mimicked a planet around the pulsar. Hundreds of sigmas crumbled to dust just before Lyne was to present his findings.

Lyne gave a presentation anyhow—a retraction. "It was an extremely difficult time," he says. "It was a large audience of extremely eminent astronomers and scientists." But at the end of his presentation, the audience broke out into a long, loud round of applause. "Here I was, with the biggest blunder of my life and …" Lyne pauses, gathering himself. "But I think that many people have nearly done such things themselves."

Lyne's reputation didn't suffer; other planet hunters weren't quite so lucky. Peter Van de Kamp of Swarthmore College in Pennsylvania will always be known as the one who found the planet around Barnard's Star. It was a planet that made it into the textbooks, even though it didn't exist.

According to George Gatewood, an astronomer at Allegheny University in Philadelphia, Van de Kamp was a victim of an equipment change. A lens assembly in the telescope had a color error that shifted redder stars with respect to their bluer counterparts. During the time that Van de Kamp observed Barnard's Star (which is red), Swarthmore replaced the assembly with one that had less color error. Barnard's Star seemed to move com-



Tantalizing. Possible sightings like this fuel CERN's hopes of winning partial credit for discovering the Higgs boson.

CREDIT: (BOTTOM) CERN

pared with the background. Years later, Swarthmore discovered a problem with the new assembly and went back to the old one. The red stars moved back into their original position. It looked as though Barnard's Star had wobbled, pulled by an imaginary planet.

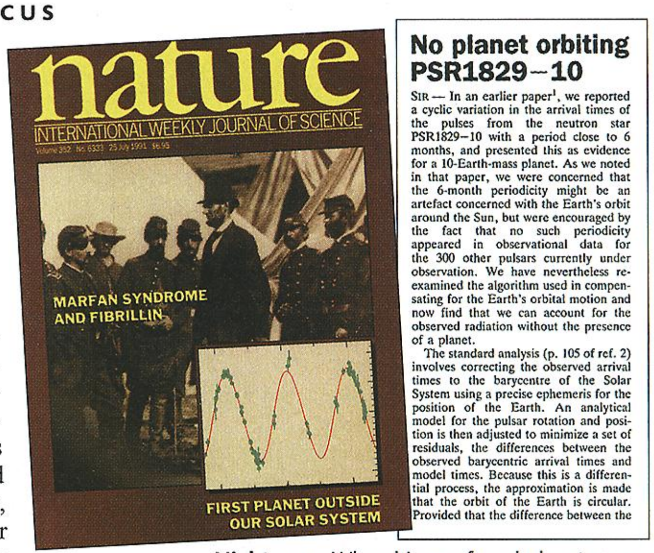
Vanishing chances

As in the heavens, so, more subtly, on Earth. Going by statistics, if physicists discovered a new five-sigma particle every day, mistaken sightings ought to turn up about once every 10,000 years. In fact, the history of high-energy physics is littered with five-sigma mirages. One was the "split A2," an unexpected double peak that, in the 1960s, seemed to signal the existence of two particles where only one was expected. "It was believed by everybody," Bahcall says. But as scientists made more measurements, the two peaks filled in, and the mysterious second particle vanished. The story replayed itself in the early 1980s, when physicists at Stanford, at DESY in Hamburg, and elsewhere found something that looked remarkably like a Higgs boson at an energy of about 8 giga electron volts (GeV), well short of the 114 GeV where CERN's current Higgs candidate lurks (*Science*, 31 August 1984, p. 912). The discovery, dubbed the zeta particle, had a five-sigma significance, but it didn't survive for long. "They kept measuring, and it disappeared," SLAC's Riordan recalls. Physicists on the zeta particle team still suffer from the memory.

A FEW "SURE THINGS" THAT WEREN'T			
Split A2	Berkeley and elsewhere, mid-1960s	5 sigma	A2 meson looked like two particles, until twin energy peaks fused
Zeta	DESY (Hamburg), Stanford University, and elsewhere, early 1980s	5 sigma	Apparent particle, thought to be the Higgs boson, died amid controversy in mid-1990s
"Axion"	GSI (Darmstadt) and elsewhere, mid-1980s	6 sigma	New heavy neutral particle vanished on closer inspection
PSR1829–10's planet	Jodrell Bank Observatory, Manchester, U.K., 1991–92	"Hundreds of sigmas"	Signs of body circling a pulsar turned out to be due to Earth's orbit

Decades of such reverses have taught experimental physicists that five-sigma rules and one-in-a-million errors are not to be taken literally. "[The statistical analysis] is based upon the assumption that you know everything and that everything is behaving as it should," says Val Fitch, who won the 1980 Nobel Prize in physics for discovering charge-parity violation in K mesons. "But after everything you think of, there can be things you don't think of. A five-sigma discovery is only five sigma if you properly account for systematics."

But if good stats alone can't guarantee a discovery's acceptance, neither do mediocre ones necessarily spell its doom. In fact, many key advances in modern physics have been accepted before passing the five-sigma test. "Neutrino mass is taken seriously, even though it's not five sigma currently," CERN's Igo-Kimenes points out—partly, Bahcall adds, because the discovery was one that physicists had long expected. "You ask, 'Does it contradict other things you already know; does it fit in with theory and experiment?' " If an observation seems to fit, scientists need less convincing to accept it, whereas extraordinary claims require extraordinary proof.



Nightmare. When his newfound planet evaporated, Andrew Lyne feared for his career.

In this regard, CERN's Higgs candidate scores high. Calculations from SLAC show that the Higgs particle should appear at an energy below 140 GeV—right in the range where the LEP team is looking. Other omens are less favorable. To get stronger confirmation of a particle's existence, physicists will often crank up their accelerator's power to see how the effects they are observing change with increasing energies. The Higgs hunters at CERN can't do that, because they are already running LEP near its maximum power to squeeze the most work out of it during its remaining weeks. Furthermore, the hints of success that CERN has reported have come from only two of LEP's four detectors. One of them sees a strong effect (three events that hint at a Higgs particle), a second sees a weak effect (one Higgs candidate), and the others see nothing. The putative Higgs particle is not behaving quite as expected: So far it appears to be sticking to one set of decay paths and ignoring other decay paths that theory says it should be taking. Moreover, LEP has been detecting three times as many of the particles as theory says it should, if the Higgs particle does indeed have the 114-GeV energy that CERN's results suggest.

Such anomalies raise the stakes for LEP's final month in two ways. On the one hand, they make it appear more likely that the collider is chasing a will-o'-the-wisp. On the other hand, if LEP now confirms those results—if its two other detectors spot Higgs candidates, or if the Higgs candidates start using other decay paths—the CERN sightings become more believable, even with few sigmas to back them up. "People will be very, very excited," says Sau Lan Wu, a CERN physicist working on one of the four detectors.

The final reason for CERN to try to beat the odds is that it has nothing to lose. If LEP doesn't find the Higgs particle, then another collider—the Tevatron at the Fermi National Accelerator Laboratory, CERN's archrival in Batavia, Illinois—probably will. By staking a claim to the Higgs particle now—and pegging it to a specific energy, 114 GeV—CERN will claim a share of the glory should Tevatron confirm the discovery, according to Riordan. "They can get a piece of it by writing an 'evidence-for' paper," he says. They might even be able to lobby for a further LEP extension and aim for a bona fide discovery, Wu hopes. And if the worst happens and the results disappear, then CERN's Higgs particle will join the ever-growing parade of ghost discoveries and phantom particles vanquished by the progress of science.

–CHARLES SEIFE