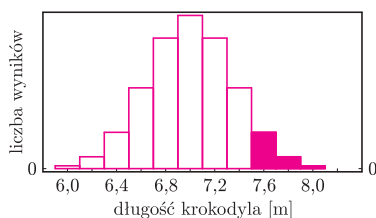
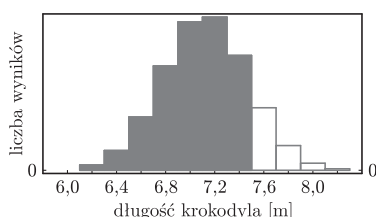


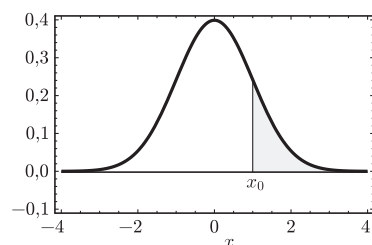
Cytaty i większość przykładów zawartych w niniejszym artykule oparte są na wykładzie, jaki wygłosił Tommaso Dorigo podczas 3rd International Conference on New Frontiers in Physics latem 2014 roku na Krecie.



Rys. 1. Rozkład wyników pomiaru długości siedmiometrowego krokodyla. W 7,6% przypadków uzyskuje się wynik co najmniej równy 7,5 m.



Rys. 2. Rozkłady wyników pomiaru długości krokodyla mierzącego 7,12 m. W 86% przypadków uzyskujemy wynik nieprzekraczający 7,5 m.



Rys. 3. Wykres gęstości rozkładu Gaussa (1). Pole pod krzywą jest równe 1; pole zacieniowanego obszaru odpowiada prawdopodobieństwu uzyskania wyniku równego co najmniej x_0 .

Magiczna liczba 0,000029%

Krzysztof TURZYŃSKI

Żeby zrozumieć coś z zachodzących w świecie procesów, należy to i owo pomierzyć. Żeby odkryć coś nowego – trzeba wykonać pomiar i uzasadnić, że nie zgadza się on z istniejącą wiedzą. Jednak wykonywanie pomiarów na granicy możliwości doświadczalnych to nie taka prosta sprawa.

Przypuśćmy, że podręcznik zoologii stwierdza autorytatywnie, iż krokodyl nie miewają więcej niż siedem metrów długości. Mamy oto przed sobą wyjątkowo dorodny okaz śpiącego krokodyla oraz dwudziestocentymetrową linijkę. Podchodzimy ostrożnie do gada i przymierzamy doń odpowiednią liczbę razy nasze narzędzie pomiarowe. Niestety, okropnie trzęsą się nam ręce, a oczy zachodzą mgłą, co wprowadza element losowości do pomiaru. Na podstawie tych wyników chcielibyśmy zdecydować, czy możemy odrzucić *hipotezę zerową* mówiącą, że mierzony krokodyl, tak jak i wszystkie inne krokodyle, ma mniej niż siedem metrów. W skrajnym przypadku siedmiometrowego krokodyla z uwagi na ów element losowy otrzymamy rozkład wyników pomiarów przedstawiony na rysunku 1. Z niezerowym prawdopodobieństwem pomiar długości siedmiometrowego krokodyla da wynik większy niż siedem metrów. Możemy jednak przyjąć lepsze kryterium: wynik przekraczający 7,5 m dostaniemy z prawdopodobieństwem jedynie 7,6%. Uznamy zatem krokodyla za ponad siedmiometrowego, gdy wynik pomiaru przekracza 7,5 m. Nie jest to rozwiązanie idealne, możemy bowiem popełnić *błąd pierwszego rodzaju* polegający na pochopnym odrzuceniu hipotezy zerowej, ale mylimy się tylko w 76 przypadkach na 1000, co nie jest w końcu takie złe.

Unikanie błędu pierwszego rodzaju ma swoją cenę. Rozważmy *hipotezę alternatywną*, że długość krokodyla jest równa 7,12 m. Analogiczny do poprzedniego rozkład wyników pomiarów jest przedstawiony na rysunku 2, gdzie zacięniowano przypadki odpowiadające wynikom pomiaru poniżej 7,5 m – czyli znaczną większość. Okazuje się, że w rozważanym przypadku nader często, bo w 86 przypadkach na 100, popełniamy *błąd drugiego rodzaju*, czyli nie odrzucamy hipotezy zerowej, chociaż poprawna jest hipoteza alternatywna. Strzegąc się przed błędem pierwszego rodzaju, łatwo zatem popełnić błąd drugiego rodzaju – i na odwrót.

Jakie zatem należy przyjąć dopuszczalne prawdopodobieństwo popełnienia błędu pierwszego rodzaju? Odpowiedź zależy od tego, co właściwie badamy. W naukach o życiu prawdopodobieństwo to ustala się często na poziomie 5%, fizyków cząstek elementarnych satysfakcjonuje zaś dopiero wartość 0,000029%. Dlaczego właśnie taka? O tym za chwilę. Najpierw musimy bowiem zapoznać się z powszechnie używanym przez fizyków słownictwem.

Rozkłady możliwych wyników pomiarów nader często przypominają kształtem *rozkład Gaussa*, którego gęstość dana jest wzorem:

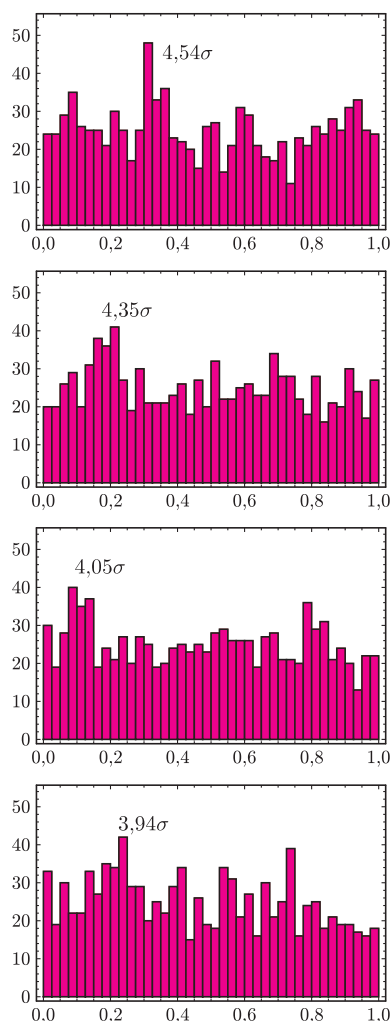
$$(1) \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Dla takiego rozkładu możemy, podobnie jak dla zagadnienia pomiaru długości krokodyla, przyjąć, że mierzona wartość x jest z zadowalającym prawdopodobieństwem większa od zera, jeśli przekracza ona pewną ustaloną wartość x_0 (rys. 3). Wartości x_0 i odpowiadające im prawdopodobieństwa zebrane są w poniższej tabeli.

x_0	1	2	3	4	5
prawdopodobieństwo	16%	2,3%	0,13%	0,0032%	0,000029%

Prawda, że ulubione prawdopodobieństwo fizyków cząstek elementarnych nie wydaje się już taką przypadkową liczbą? Ponieważ wartość $x_0 = 1$ odpowiada *odchyleniu standardowemu* rozkładu Gaussa, które oznacza się tradycyjnie przez σ , więc poszczególne prawdopodobieństwa możemy zapisać umownie jako 1σ , 2σ itd.

Zastanówmy się teraz, jak fizycy odkrywają nowe cząstki. Pisał o tym w *Delcie* 12/2013 Maciej Misiura, więc przypomnimy tutaj tylko pokrótce najważniejsze wiadomości. Wiedząc, na jakie odpowiednio stabilne cząstki poszukiwana cząstka nietrwała może się rozpaść, należy pomierzyć energie i pędy cząstek potomnych. Znając te wielkości, możemy obliczyć *masę niezmienniczą*, która odpowiada masie rozpadającej się cząstki. Niestety, z uwagi na losowe efekty w detektorze energie



Rys. 4. Wybrane histogramy przedstawiające rozkłady liczb losowanych z przedziału (0, 1). Przy „garbach” podano obliczone na podstawie rozkładu Poissona prawdopodobieństwo uzyskania tak dużej sumy wyników w trzech przedziałach w pojedynczym losowaniu. Po uwzględnieniu liczby niezależnych trójek sąsiadujących przedziałów w histogramie prawdopodobieństwa te zmniejszają się odpowiednio do $3,97\sigma$, $3,74\sigma$, $3,39\sigma$ i $3,27\sigma$.

Leon Lederman (ur. 1922 r.) jest jednym z najwybitniejszych fizyków doświadczalnych zajmujących się cząstkami elementarnymi. Współtwórca potwórka językowego *boska cząstka* na określenie cząstki Higgsa (z D. Teresim), u szczytu swej kariery dokonywał ważnych odkryć w takim tempie, że pominięte przezeń rezultaty wystarczyłyby pewnie na kilka Nagród Nobla. Doskonałym przykładem może być niedocenie wagi procesów słabych z udziałem tzw. prądów neutralnych: Lederman te reakcje zarejestrowane w swoim eksperymencie nazwał dosadnie „gównionami” (ang. *crapon*).

i pędy znamy ze skończoną dokładnością, zatem zamiast otrzymywać za każdym razem tę samą wartość masy niezmienniczej, otrzymamy pewien jej rozkład. Co gorsza, detektor może zarejestrować także *tło*, czyli cząstki powstające w wyniku innych procesów niż rozpad poszukiwanej cząstki. Poszukiwanie nowych, nietrwałych cząstek sprowadza się zatem przede wszystkim do poszukiwania „garbów” w rozkładach odpowiednio wyznaczonej masy niezmienniczej przy uwzględnieniu tła.

Niestety, także w przypadku tła obecny jest nasz nieszczęsny element losowy. Może to znacząco utrudnić decyzję, czy obserwowany w badanym rozkładzie masy niezmienniczej „garb” jest efektem losowym, czy też przejawem nowego, poszukiwanego procesu fizycznego polegającego na rozpadzie nieznannej cząstki. Łatwo to pokazać na następującym przykładzie. Losujemy 1000 liczb z przedziału (0, 1), a następnie grupujemy je w 40 równych długości przedziałach. Powtarzamy tę procedurę 2000 razy i wybieramy w jakiś sposób rozkłady z najbardziej wyraźnymi „garbami”. Jeden z możliwych wyników tej zabawy przedstawiony jest na rysunku 4. Jak zdecydować, czy obecne na rysunkach „garby” pochodzą od losowych wartości tła? Pół biedy, jeżeli wiemy, spośród ilu rozkładów zostały wybrane te „garbate”. Jak jednak podjąć taką decyzję, jeśli to kolega pokazuje nam takie obrazki, a my nie wiemy, jak bardzo się starał, żeby je wyprodukować? Nie sztuką jest nawoływać: „Bądźmy konserwatywni i nie popełniamy błędów pierwszego rodzaju!”, jak jednak pocieszyć naukowca, któremu w ten sposób uciekło odkrycie i – być może – Nagroda Nobla?

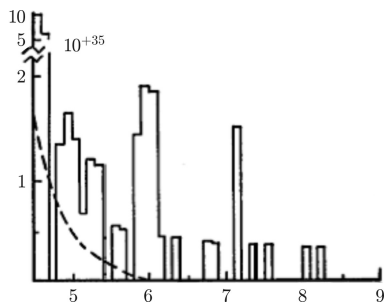
Problemem tym zajął się w 1968 roku Arthur Rosenfeld. W poświęconym odkrywaniu nowych, egzotycznych cząstek artykule pisał, że każdego roku masę niezmienniczą wyznacza się około 35 milionów razy. Rzut oka na czasopiśmie naukowe pozwala stwierdzić, że typowy histogram dla takich danych opiera się na około 2500 liczbach, co oznacza, że rocznie sporządza się około 15 tysięcy takich histogramów. Na tej podstawie Rosenfeld oszacował, że każdego roku analiza tych histogramów powinna dawać od 10 do 100 pozornych „garbów”, błędnie interpretowanych jako nowe cząstki (większa z tych liczb uwzględnia fakt, że gdy badacz uwierzy już w odkrycie zawarte w przygotowanych przezeń danych, będzie dane te obrabiał tak, by odkrycie to stało się możliwie bezdyskusyjne). Porównanie z publikowanymi w literaturze naukowej doniesieniami o nowych cząstkach okazało się zgodne z tym zgrubnym oszacowaniem. Z tego względu Rosenfeld podsumował swój artykuł następująco.

Dla teoretyka bądź fenomenologa moral jest prosty: trzeba czekać na efekty na poziomie 5σ . Może to być trudniejsze dla zespołu doświadczalnego, który zainwestował rok pracy i jakiś milion dolarów, [...] można więc opublikować takie dane, [...] powinno być jednak jasne, że każdy garb poniżej poziomu 5σ wskazuje na konieczność powtórzenia doświadczenia.

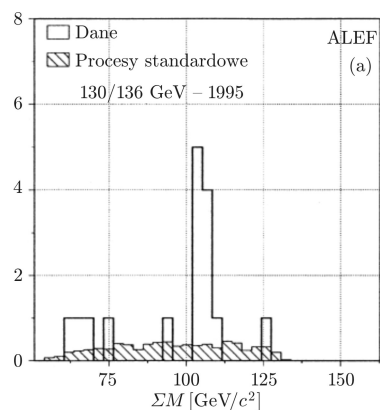
Apel Rosenberga nie od razu jednak został potraktowany poważnie. Można się o tym przekonać, przyglądając się niektórym odkryciom w fizyce cząstek elementarnych z ostatnich 40 lat.

W 1974 roku odkryto czarmonium, J/ψ , cząstkę złożoną z kwarka powabnego i jego antycząstki. „Garby” były w tym przypadku tak doskonale widoczne, że nikt się nie przejmował porządną analizą statystyczną. Podobna sytuacja miała miejsce kilka lat później przy odkrywaniu taonu.

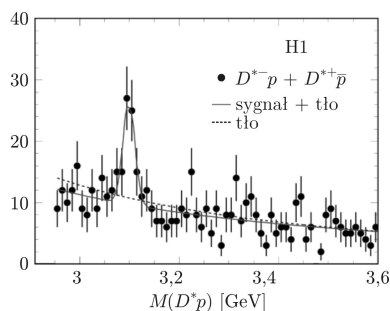
W 1976 roku grupa fizyków z Fermilabu, kierowana przez Leona Ledermana, odkrywcę neutrino mionowego i późniejszego laureata Nagrody Nobla, doniosła o zaobserwowaniu nowej cząstki o masie równej około siedmiu masom protonu ($6,5 \text{ GeV}/c^2$) i zaproponowała dla niej oznaczenie Υ . Rozkład liczby zliczeń par elektron-pozyton o różnych masach niezmienniczych przedstawiony jest na rysunku 5; prawdopodobieństwo, że taki rezultat wynika po prostu z losowych własności tła został oszacowany na 2% (około 2σ). W miarę przybywania danych okazało się jednak, że widoczny na wykresie „garb” po prostu zanikł po uwzględnieniu świeższych wyników, a więc hipoteza zerowa (brak nowej cząstki) została w tym przypadku odrzucona nieco zbyt pochopnie. Złośliwość fizyków sprawiła, że w annałach historii nauki był, którego istnienie zaanonsował Lederman *et consortes*, został uwieczniony jako pomyłka tego fizyka i wynik ten określa się dziś hasłem *Oops-Leon*. Oznaczenie Υ przydało się zaś dla bottomonium, cząstki o masie około dziesięciu mas protonu odkrytej przez ten sam zespół rok później. Nauczni przykrym doświadczeniem badacze po zaobserwowaniu „garbu”, który mógł być zgodny z hipotezą braku nowej cząstki jedynie na poziomie przekraczającym 3σ , cierpliwie poczekali, czy i ten nie „rozplynie” się w później uzyskanych danych.



Rys. 5. Rozkład tempa produkcji par elektron-pozyton w zależności od masy niezmienniczej, który doprowadził do ogłoszenia odkrycia cząstki Υ . Źródło: Horn i wsp., Phys. Rev. Lett. 36 (1976) 1236.



Rys. 6. Rozkład masy niezmienniczej w przypadkach z czterema wiązkami (dżetami) cząstek uzyskany przez zespół ALEPH. Zakresowany histogram odpowiada oszacowanemu poziomowi tła. Na podstawie: Buskalic i wsp., Z. Phys. C71 (1996) 179.



Rys. 7. Uzyskany przez zespół H1 rozkład masy niezmienniczej, na podstawie którego wysunięto hipotezę o istnieniu pentakwarków. Na podstawie: Aktas i wsp., Phys. Lett. B588 (2004) 17.

W styczniu 1983 roku ogłoszono odkrycie w CERN-ie cząstki W . Artykuł naukowy, w którym to uczyniono, opierał się na sześciu zaobserwowanych przypadkach-kandydatach, nie podawał prawdopodobieństwa, z jakim odrzuca się hipotezę zerową, zauważał jedynie, że prawdopodobieństwo zaobserwowania takiej sytuacji wskutek losowych własności tła jest niezwykle małe. Tak samo rzecz się miała z ogłoszeniem odkrycia cząstki Z w maju tego samego roku, na podstawie czterech zaobserwowanych reakcji.

Reguła Rosenberga znalazła bardziej rygorystyczne zastosowanie dopiero w nieco nowszych czasach. W 1994 roku badacze z zespołu CDF w Fermilabie widzieli w rozkładach masy niezmienniczej „garb” w miejscu odpowiadającym temu, gdzie spodziewano się znaleźć kwark top. Ponieważ jednak nie byli w stanie odrzucić hipotezy zerowej z prawdopodobieństwem przekraczającym poziom 3σ , w pierwszych swoich artykułach pisali o argumentach (ang. *evidence*) na rzecz istnienia kwarka top. Dopiero po roku, kiedy zarówno CDF, jak i D0, dwa zespoły z Fermilabu, badające ten wycinek rzeczywistości, uzyskały odrzucenie hipotezy zerowej na poziomie przekraczającym 5σ , zdecydowano się mówić o odkryciu nowej cząstki – i to rok 1995 przyjmuje się za datę odkrycia kwarka top.

Cząstka Higgosa także była odkrywana dwa razy. Autor tego artykułu doskonale pamięta napięcie, z jakim w grudniu 2011 roku śledził transmisję na żywo z CERN-u, gdzie ogłoszono zaobserwowanie w trzech rodzajach rozkładów masy niezmienniczej „garbu” na poziomie 3σ w miejscu odpowiadającym właśnie cząstce Higgosa o masie około $125 \text{ GeV}/c^2$. Wydarzenie to zapoczątkowało prawdziwy wysyp prac naukowych analizujących własności tej cząstki ze wszystkich możliwych punktów widzenia. CERN-owska konferencja prasowa z 4 lipca 2012 roku, na której ogłoszono przekroczenie poziomu 5σ i której datę na ogół przyjmuje się za oficjalną „chwilę” odkrycia cząstki Higgosa, w świadomości naukowców zmieniła już niewiele.

Nie wolno też zapominać o blamażach w stylu *Oops-Leon*, przed jakimi reguła Rosenberga pozwoliła środowisku naukowe uchronić. W 1996 roku badacze z zespołu badawczego używającego detektora ALEPH przy akceleratorze LEP w CERN-ie, poszukując cząstki Higgosa, uzyskali rozkład masy niezmienniczej przedstawiony na rysunku 6. W tym przypadku hipotezę zerową można było odrzucić z prawdopodobieństwem 0,01% ($3,7\sigma$), jednak dalsze badania nie potwierdziły zaobserwowania nowej cząstki.

Nie zawsze jednak wystarczy ślepo stosować regułę Rosenberga, by nie popełnić błędu pierwszego rodzaju. W 2004 roku zespół badawczy używający detektora H1 przy akceleratorze Hera pod Hamburgiem uzyskał przedstawiony na rysunku 7 rozkład masy niezmienniczej, kuszący wręcz, by ogłosić światu istnienie pentakwarka, cząstki złożonej z pięciu kwarków i antykwarków – tyle że jej istnienia nikt się właściwie nie spodziewał. Odrzucenie hipotezy zerowej było w tym przypadku błędem z prawdopodobieństwem zaledwie $3 \cdot 10^{-10}$ (ponad 6σ), jednak analizujący te dane fizycy, najwyraźniej uznając, że rewolucyjne doniesienia wymagają nadzwyczajnie mocnych dowodów, skromnie użyli frazy „argumenty na rzecz istnienia” w tytule swego artykułu. Jak się okazało, słusznie, bo istnienia pentakwarków potwierdzić się nie udało.

Niekiedy za nazbyt łatwe odrzucanie hipotezy zerowej odpowiadają nie efekty losowe, ale te związane ze sposobem przeprowadzenia doświadczenia (systematyczne). Zespół badawczy, używający detektora OPERA w Gran Sasso we Włoszech rejestrującego neutrino z wiązki wyprodukowanej w CERN-ie, ogłosił w 2011 roku, że z prawdopodobieństwem na poziomie 6σ należy odrzucić hipotezę, że neutrino są cząstkami poruszającymi się co najwyżej z prędkością światła. Jak się okazało w kolejnym roku, owo 6σ nie miało nic wspólnego z rzeczywistością, a cały efekt spowodowany był opóźnieniem rejestracji neutrin wskutek luźno wetkniętego kabla.

Czy wszystko to oznacza, że nie można wierzyć nawet „prawdom naukowym”? Nic podobnego! Jak wynika z przytoczonych wyżej przykładów, pierwsza reakcja badacza uzyskującego nowy, zaskakujący wynik lub o takim wyniku się dowiadującego polega na próbie zgromadzenia większej liczby danych, niezależnego powtórzenia doświadczenia lub zastanowienia się, czy podczas doświadczenia nie popełniono jakichś błędów. W ten sposób nauka nieustannie się oczyszcza z błędnych wniosków, także i tych, które są w stanie przejść przez sito zasady Rosenberga. I właśnie dlatego nauka jest nam w stanie co nieco powiedzieć o świecie.