

## **Empiryczne wprowadzenie do statystyki**

Tadeusz Wibig

Punsiri Dam-o

Uniwersytet Łódzki, Katedra Modelowania Procesów Nauczania

*Streszczenie:* W pracy przedstawiamy nasze doświadczenia z prowadzonych na Wydziale Fizyki i Informatyki Stosowanej Uniwersytetu Łódzkiego lekcji doświadczalnych wprowadzających do statystyki młodzież pierwszych klas liceów biorących udział w programie „klas pod Patronatem”. Nowa, oryginalna metoda nauczania podstawowych elementów statystycznego opracowywania danych pomiarowych, wprowadza pojęcia niepewności pomiarowej, niepewności średniej, średniej ważonej. Metoda nie wymaga żadnych specjalnych umiejętności, czy wiedzy matematycznej. Użycie na zajęciach instrumentu, który nazwaliśmy „takoyaki” i który wraz z zestawem eksperymentalnym kolorowych kulek stanowi istotny element uatrakcyjnienia zasadniczo nieciekawej lekcji polegającej na powtarzaniu wiele (set) razy prób Bernoullego. Proponowana metoda sprawdza się już od kilku lat realizacji projektu.

### **Wstęp**

Formalne podejście do niepewności pomiarowych i ich statystycznej interpretacji opisane jest w dokumencie Międzynarodowej Organizacji Standaryzacji (IOS) pod nazwą „Guide to the Expression of Uncertainty in Measurement” (znanego pod skróconą nazwą GUM) [1]. Ze względu na wagę instytucji zalecany jest on z oczywistych względów do stosowania na zaawansowanych poziomach edukacji [2]. Oficjalny dokument, czy to w wersji IOS, czy Narodowego Instytutu Standaryzacji i Technologii (NIST), Komitetu Europejskiego (EURACHEM/CITAS) czy brytyjskiego (UKAS) [3,4] nie nadaje się jednak zupełnie do użytku w procesie nauczania, a szczególnie, gdy chodzi o wstępne jego poziomy. Aby mógł być zastosowany niezbędne jest poczynienie w nim istotnych zmiany, co najmniej jeśli chodzi o formę przekazu. Jest to problem w zasadzie techniczny i gdyby był jedynym do pokonania, nie byłoby w ogóle problemu.

Instrukcja GUM z istoty swojej nie jest przeznaczona do nauczania kogośkolwiek, a w szczególności uczniów z zerową, lub niemal zerową, wiedzą na temat statystyki. W najlepszym razie jej adresatami są ci, którzy już wiedzą, o co chodzi, albo też ci, których wiedza ta wcale nie interesuje, a poszukują prostego przepisu postępowania z wynikami konkretnych pomiarów. Postępowanie zgodnie z zasadami GUM jest w sumie nieskomplikowane i numerycznie proste. Dotyczy ono jednak tylko standardowych przypadków, a właśnie szkolne eksperymenty fizyczne z reguły są proste i nikt (na początku) nie wymaga od uczniów rozumienia zasad statystycznej analizy danych. Jednocześnie właściwe użycie algorytmów GUM przy elementarnej zdolności wykonania odpowiednich rachunków, jak pokazują badania [2], nie skutkuje żadną głębszą refleksją natury

statystycznej. Ograniczanie się więc do przekazania uczniom zestawu, słusznych skądinąd formuł, nie prowadzi do zrozumienia czegokolwiek, a więc tym bardziej do rozwoju umożliwiającego kiedyś, w przyszłości, choćby potencjalnie, rozwiązywanie problemów niestandardowych.

Poznanie rzeczywistości opiera się dziś na metodzie zwanej naukową, a podstawą metody naukowej jest zadawanie Naturze pytań i notowanie jej odpowiedzi. Zwykle nazywa się to „eksperymentowaniem”. Uznając tę metodę, wydaje się, że w procesie nauczania powinniśmy starać się zachęcać uczniów, aby zaczęli świadomie eksperymentować najwcześniej, jak to jest możliwe, najlepiej jeszcze w szkole podstawowej. Z drugiej strony możliwe jest też poczekać do chwili, gdy uczniowie nabędą umiejętności posługiwania się metodami GUM i gdy pojmą sens formuł wyrażających niepewności pomiarowe i nauczą się podawać wynik doświadczenia w sposób jedyny słuszny, pełny i właściwy (GUM). W takim jednak przypadku moglibyśmy zacząć eksperymentować dopiero w laboratoriach uniwersytetów. Postawieni przed taką alternatywą nauczyciele gimnazjów i liceów muszą wybrać własny „złoty środek”. W jego znalezieniu są oni jednak pozostawieni samym sobie. Poniżej podajemy metodę umożliwiającą właściwe (zgodnie z GUM!) zinterpretowanie wyników doświadczeń na poziomie „przedobliczeniowym” [5], na początku naukowej edukacji.

### **Początek lekcji: niepewność pomiarowa**

Jednym z najważniejszych problemów w naukach doświadczalnych jest określenie niepewności pomiarowej pojedynczego pomiaru. Powtarzanie pomiaru „w dokładnie takich samych warunkach” może dać nam i najczęściej daje za każdym razem inny wynik. Rozrzut tych wyników nazywa się czasem „odchyleniem standardowym”, ale ponieważ pojęcie to jest precyzyjnie zdefiniowane (w GUM), by ustrzec się konfliktu z przyjętą terminologią nazywać go będziemy chwilowo „niepewnością pojedynczego pomiaru” (za: [6]).

Dla porządku też w dalszej części tej pracy pominiemy problemy z niepewnościami typu B [4] i skupimy się na niepewnościach mających charakter czysto statystyczny.

Punktem wyjścia naszych działań na lekcji w klasie „pod Patronatem” jest uświadomienie uczniom, że we wszystkich, albo, co najmniej w lwiej części, wszystkich istotnych pomiarów ich wynik ma w jakimś stopniu charakter statystyczny. W terminologii GUM: niepewność typu A w pomiarze obecna jest prawie zawsze, nawet, gdy na pierwszy rzut oka jej nie zauważamy, albo nie spodziewamy się jej w ogóle. Jest to dość proste i możemy powołać się na wiele przykładów „z życia wziętych”. W klasie możemy w tym celu wykorzystać proste, wstępne doświadczenie „spadku swobodnego” mierząc czas opadania dowolnego przedmiotu podrzuconego do góry pod sufit od chwili, gdy osiągnie położenie najwyższe, do momentu uderzenia w podłogę. Uczniowie mogą zmierzyć ten czas używając stoperów wmontowanych w ich własne telefony komór-

kowe. Na ogół z jednego rzutu otrzymujemy około 20 wyników, które po zapisaniu na tablicy (Rys. 1) pokazują rozrzut zmierzonych wartości, czyli właśnie „niepewność pomiarową pojedynczego pomiaru”. Wyniki te należy zachować, gdyż posłużą nam jeszcze w dalszych częściach lekcji.



Rysunek 1. Wyniki doświadczenia ze spadkiem swobodnym (w setnych częściach sekundy) zmierzone na lekcji.

Rozważania nasze nad statystyką zaczynamy od prostego pytania: „Jaki jest wynik naszego eksperymentu ze swobodnym spadkiem?”, „Jak długo spadało ciało?”. Zazwyczaj kilku co sprytniejszych uczniów próbuje odpowiedzieć na to pytanie wyciągając kalkulatorki (albo wykorzystując coraz bardziej rozbudowane telefony komórkowe) i próbując obliczyć wartość średnią, a precyzyjnie mówiąc średnią arytmetyczną.

Oczywiście jest to, w większości przypadków, odpowiedź właściwa. Posługując się całą wiedzą GUM nic ponad to powiedzieć nie można. Ale, jeśli spytać tego „najsprytniejszego” ucznia, dlaczego uważa on (albo ona!), że właśnie arytmetyczna średnia jest tym, co podać należy w odpowiedzi na postawione pytanie, najczęściej jasnej odpowiedzi się nie uzyska. Jeśli drążyć temat podając informacje o innych rodzajach średniej: średniej geometrycznej, średniej harmonicznej i w końcu o nieskończonej możliwości ich zdefiniowania

(przez tak zwaną p-miarę) „oczywista” odpowiedź uzyskana na początku staje się coraz mniej oczywista.

Zostawiając ten temat do dalszych rozważań, a można w tym miejscu zapowiedzieć podanie prawidłowej odpowiedzi wraz z uzasadnieniem za godzinę mniej więcej, możemy dalej spytać, jak uczniowie określiliby to, co potocznie nazywamy rozrzutem wyników, jaki widać na tablicy (Rys. 1) wyraźnie. Zwykle na tak postawione pytanie nie udaje się uzyskać żadnej „oczywistej” odpowiedzi. Poprawną odpowiedzią, wydają się jest w tym momencie: „nie wiem”.

Okazuje się, że, jak pokazują badania [7, 8], nim uczniowie dowiedzą się, jak to jest „naprawdę” (chodzi oczywiście o „odchylenie standardowe”), mają oni już w swoim umyśle intuicyjną ideę rozrzutu, o którym mowa. Jest i ją trudno wyrazić werbalnie, a jeszcze bardziej precyzyjnie zdefiniować ([9] i podane tam referencje). Pojęcie rozrzutu/dyspersji/rozmycia/odchylenia nie jest ani oczywiste, ani też jednoznaczne. W pracy [10] przeanalizowano w tym względzie różne hiszpańskojęzyczne podręczniki dla młodzieży w wieku lat 15-16 i naliczono niemal 20 możliwości jego określenia (a „odchylenie standardowe” było tylko jedną z tych możliwości!).

Metoda GUM definiuje pojęcie „rozrzutu” w paragrafie 2.4 w IOS/NIST „Guide...”, jako wspomniane już „odchylenie standardowe” równe pierwiastkowi kwadratowemu z wariancji określonej zmiennej losowej. Aby pojąć tę definicję trzeba posiadać, co najmniej elementarną wiedzę z rachunku prawdopodobieństwa, a pojęcie gęstości prawdopodobieństwa niezbędne dla zrozumienia definicji słowa „wariancja” wymaga umiejętności różniczkowania i całkowania. Jest to zdecydowanie zbyt wiele dla przeciętnych i dla większości ponadprzeciętnie zdolnych uczniów szkół średnich. Taki sposób definiowania wariancji, jest dobry dla jedynie studentów wyższych uczelni i to, na ogół, nie na pierwszym roku.

W badaniach Kazaka i Confreya [8] zostało jednak pokazane, że już nawet dziewięcioletnie dzieci potrafią pojąć i stosować w praktyce naturalne dla nich pojęcie rozrzutu (wariancji?) używając w tym celu graficznej reprezentacji danych, którą fizycy nazywają histogramem, a statystycy szeregiem rozdzielczym (jedną z metod analizowanych w [10]). Proponujemy na lekcji rozwijać to właśnie, naturalne podejście.

### **Histogram i pojęcie miary rozrzutu**

Przekazanie uczniom pojęcia rozrzutu wyników pomiarowych przy wielokrotnym powtarzaniu tego samego eksperymentu w sposób teoretycznie właściwy (GUM!) jest zadaniem trudnym. W pewnych szczególnych przypadkach „niepewność pomiarowa pojedynczego pomiaru” może być przeanalizowana teoretycznie, a najlepszym na to przykładem jest „problemie pijanego marynarza” (drunken sailor problem).

W przypadku jednowymiarowym był on omówiony w interesującym nas aspekcie ponad 50 lat temu w „Feynmana wykładach z fizyki” [11] i ponownie przywołany przez Błasiaka [6] trzydzieści lat później. Wspomniany marynarz w przypadku jednowymiarowym może zrobić krok w lewo, lub w prawo, przy czym jego świadomość na tyle odbiega od trzeźwej rzeczywistości, że każdy kolejny krok nie zależy od kroku poprzedniego, co dla prostoty możemy symulować na trzeźwo rzucając symetryczną monetę przed każdym krokiem i od wyniku (orzeł/reszka) uzależnić jego kierunek. Jednowymiarowy pijany marynarz „średnio rzecz biorąc” na pewno z każdym krokiem oddala się od punktu startu. Okazuje się, że „średni dystans” na jaki się oddalił rośnie jak pierwiastek kwadratowy z czasu pijanego spaceru. Bardziej realistyczny obrazek dwuwymiarowego pijanego marynarza zainteresował swego czasu Karla Pearsona (wynalazcę metody minimalizacji chi kwadrat!). W roku 1905, roku cudownym dla fizyki, Pearson opublikował w czasopiśmie „Nature” [12] krótki tekst z zapytaniem o to, jak przemieszczał się będzie pijany marynarz w płaszczyźnie. Nie musiał długo czekać na odpowiedź, bo już w numerze „Nature” z kolejnego miesiąca odpowiedział mu Lord Rayleigh. Odpowiedź była jedynie „asymptotyczna”, podana dla dostatecznie długich czasów, ale zadowoliła Pearsona na tyle, że w kolejnym numerze „Nature”, po tygodniu, podziękował Lordowi. Rayleigh podawał nie tylko formułę z pierwiastkiem z czasu, ale i wzór na gęstość prawdopodobieństwa tego, że marynarz dojdzie na zadaną odległość. Był to oczywiście rozkład normalny zwany rozkładem Gaussa. Odpowiedź Pearsona zawierała błyskotliwe podsumowanie wszystkich tych skomplikowanych matematycznych rozważań: „...wniosek, jaki można wyciągnąć z rozwiązania Lorda Rayleigha jest taki, że najbardziej prawdopodobnym miejscem znalezienia pijaka, jeśli ten jest jeszcze w stanie w ogóle utrzymać się na nogach, jest punkt bliski temu, z jakiego wyruszył”.

Zostawiając nietrzeźwego marynarza z jego własnymi, specyficznymi problemami, musimy stwierdzić, że rozwiązanie Rayleigha nie jest naprawdę ani proste, ani niestety ogólne, jeśli chodzi o opisanie niepewności (rozrzutu) wyników pomiarów. Studiowanie problemu pijanego marynarza wymaga głębokiej znajomości rachunku prawdopodobieństwa i, być może, poza samym tematem nie budzi szczególnego zainteresowania uczniów.

Konwencjonalny sposób wprowadzenia niepewności pomiarowych w całkowicie poprawny sposób podany jest choćby w pracy [15]. Problem niepewności pojedynczego pomiaru, jego rozrzutu, zdefiniowany z użyciem właściwych narzędzi matematycznych może być, z użyciem tych narzędzi, rozwiązany dając tym samym uczniom zarówno potrzebne definicje, jak i rozumienie analizowanych zjawisk. Jedyne kłopot, jak już wspominaliśmy, polega na tym, że uczniowie muszą być wcześniej tych narzędzi nauczeni, a jest to poziom co najmniej akademicki, o którym w szkole nie mamy co marzyć,

a w szczególności w stosunku do dzieci, które rozpoczynają swą edukację przyrodniczą.

Niekonwencjonalny, empiryczny sposób wprowadzenia niepewności zakłada minimalne zastosowanie metod matematycznych, a zamiast nich użycie zestawu eksperymentalnego, który nazywać będziemy z japońska „takoyaki”, o czym za chwilę. Przez ostatnie pięć lat przeprowadziliśmy zajęcia z ponad 30 klasami dzieci w wieku od lat 14 do 19 (studentami kierunku „Leśnictwa” na Uniwersytecie Łódzkim), zbierając doświadczenia, na mocy których jesteśmy w stanie stwierdzić, że proponowana metoda uczenia niepewności pomiarowych sprawdza się w praktyce. Możemy powiedzieć też, że metoda ta jest efektywna. Jedne zajęcia trwają przeciętnie dwie godziny, a zawierają one jeszcze mały wykład, jako historyczne wprowadzenie i kończą się podsumowaniem pokazującym istotę przeprowadzonych działań i wyprowadzenia odpowiednich wzorów w precyzyjnej matematycznej formie. Po takich zajęciach uczniowie mogą dodatkowo opracować otrzymane przez siebie wyniki i podsumować je we własnym raporcie. Ta część potencjalnej aktywności uczniowskiej pozostawiona jest do poprowadzenia nauczycielowi w szkole, na przykład na zajęciach koła fizycznego.

W zajęciach pierwszorzędą rolę odgrywa nauczanie uczniów sporządzania histogramów. Po podaniu właściwej definicji i algorytmu postępowania uczniowie przystępują do sporządzenia histogramu zmierzonych wartości czasu spadku zapisanych na tablicy Rys. 1.

Najistotniejszym elementem wspomnianego teoretycznego wprowadzenia w temat jest przekonanie uczniów, że istnieje asymptotyczna forma histogramu, jaki sporządzili i że jest nią krzywa dzwonowa odkryta przez Gaussa około 200 lat temu.

Kolejnym punktem jest pokazanie, w jaki sposób uniwersalna krzywa dzwonowa może być, a właściwie powinna być dopasowywana do realnego histogramu przy skończonej statystyce, takiego jak sporządzony właśnie w doświadczeniu ze spadaniem. Oczywiście można wprowadzić zmierzone dane (czasy spadania) do jakiegokolwiek komputerowego arkusza kalkulacyjnego, na przykład Open Office Calc, czy Microsoft Office Excel. Nie stanowi dla młodzieży szczególnego problemu, tak jak łatwe jest potem uruchomienie opcji „histogram”, czy „częstość”, a następnie „fit the best Gaussian”, „rozkład normalny”. Wynik takich operacji będzie oczywiście „doskonały”, a samo dopasowanie krzywej w magiczny sposób wprowadzi na wykres „najlepszą” linię i dla większości z dzieci, tak samo zresztą, jak i dla prawie wszystkich dorosłych użytkowników komputerów, linia ta traktowana będzie jako prawda absolutna, niepodważalna, w najmniejszym stopniu nie podlegająca dyskusji ani wątpliwości żadnej. Do kwestii dopasowywania numerycznego (komputerowego) wrócimy jeszcze niżej.

Bardzo istotne jest uczenie przyszłych eksperymentatorów od najmłodszych lat, w jaki sposób przeprowadzić ciągłą, gładką i symetryczną krzywą o dzwonowa-

tym kształcie przez otrzymany histogram tak, aby opisywała go jak najściślej, „najlepiej”, jak się tylko da. Idea takiego dopasowywania nie jest wcale trudna do pojęcia, o ile oczywiście opuści się cały skomplikowany aparat matematyczny. Rzeczą ludzką jest popełnianie błędów, takich jak kreślenie łamanej poprzez liczebności kolejnych klas histogramu, co daje rozwiązanie nieładkie i na ogół niesymetryczne, czy nadmierna waga przywiązywana do wyników pojedynczych trafień dalekich od spodziewanej wartości, co daje krzywe zbyt szerokie, przeszacowujące wartości pośrednie, albo przykładanie nadmiernego znaczenia do liczebności klasy najliczniejszej, a to prowadzi znów do przeszacowania polegającego na rysowaniu linii w zasadzie ponad wszystkimi innymi słupkami histogramu i tak dalej, i tak dalej...

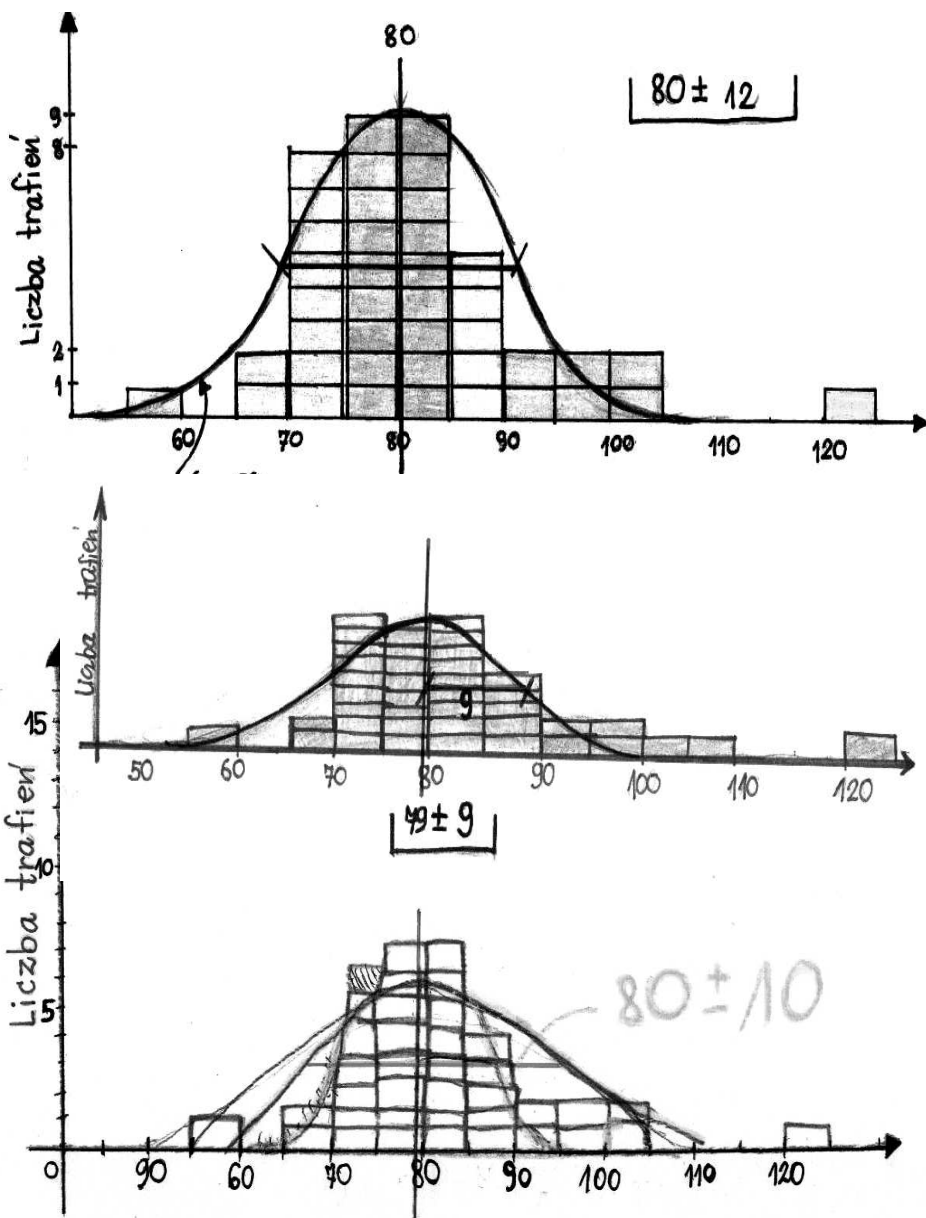
Wskazanie uczniom na te błędy i wyjaśnienie, jak należy krzywą Gaussa dopasować do histogramu, aby pozostać w zgodzie z zasadami, jest zadaniem trudnym, ale wartym włożonego wysiłku, gdyż umiejętność ta raz nabyta procentować będzie w przyszłości i niełatwo ulegnie zapomnieniu, jak jazda na rowerze, o czym mogliśmy się przekonać na kilku przykładach. Nauczenie tego wymaga udziału i pomocy osób wiedzących, o co chodzi (w programie realizowanym na Uniwersytecie Łódzkim pomagają w tym doktoranci Wydziału Fizyki i Informatyki Stosowanej), ale mogą to być też nauczyciele znający niuanse dopasowywania. A poza tym uczenie się na własnych błędach, poza tym, że efektywne, zapewnia to, że u dzieci pojawia się naturalne pojęcie dopasowania, jego sensu i jakości. Ważne jest, aby, nim wykonamy następny krok ku zrozumieniu statystyki, uczniowie potrafili możliwie poprawnie kreślić linie, zwracając uwagę na symetrię, normalizację i kształt krzywej Gaussa. Przykłady dopasowań zrobione przez uczniów pokazuje Rys. 2.

Aby wyciągać wnioski ilościowe i wykorzystać właściwie sporządzone histogramy, musimy jeszcze zdefiniować kilka bardzo istotnych pojęć. Musimy też zdefiniować niepewność. W naszym przypadku nie chodzi o jakąś formułę, wzór do zapamiętania, który prędzej, czy później zostanie przecież zapomniany, ale o całą procedurę, która sama w sobie nie może być specjalnie skomplikowana.

Rozrzut, czyli inaczej mówiąc niepewność, pojedynczego pomiaru definiujemy jako szerokość właściwie dopasowanej, co ustaliliśmy wcześniej, krzywej Gaussa – precyzyjnie jako **„połowę szerokości w połowie wysokości”**, co jest dość łatwe do zapamiętania po kilkukrotnym powtórzeniu.

Mając nakreślone najlepsze dopasowanie łatwo jest znaleźć maksimum krzywej dzwonowej. Jest to po prostu najwyższy jej punkt. Nie wymaga to żadnej wcześniejszej wiedzy. Dalej przez ten punkt poprowadzimy linię pionową. Linia ta przecnie oś „ $x$ ” w punkcie, który możemy nazwać wartością najbardziej prawdopodobną, a zatem jest to punkt, który odpowiada na pytanie: „jaki jest wynik naszego pomiaru”. Odcinek od maksimum, do punktu na osi „ $x$ ”, wysokość krzywej Gaussa, można bez specjalnych narzędzi, podzielić na pół. Przez punkt podziału możemy oczywiście narysować linię poziomą i znaleźć szerokość krzywej Gaussa na tym poziomie: długość odcinka linii poziomej wyciętego krzywą Gaussa. Połowa długości te-

go odcinka wyrażona w odpowiednich jednostkach (takich jakimi mierzona jest wartości zmiennej „x” zaznaczone na osi poziomej) jest właśnie tym co nazywać będziemy „**połową szerokości w połowie wysokości**”, zdefiniowaną przez nas właśnie miarą rozrzutu pojedynczych pomiarów.



Rysunek 2. Trzy przykłady histogramów sporządzonych przez uczniów wraz z dopasowanymi przez nich „najlepszymi” krzywymi Gaussa.

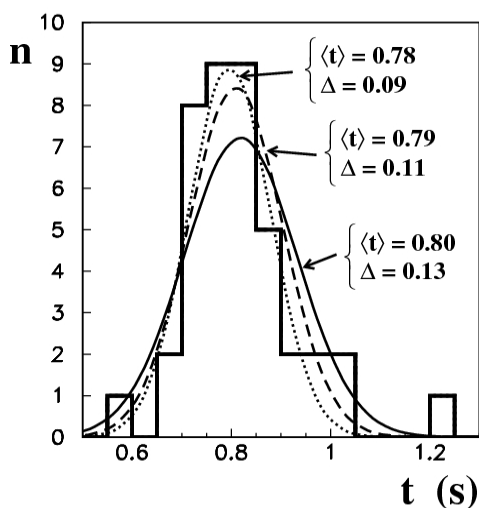


Nie jest to oczywiście znane powszechnie odchylenie standardowe nazywane niepewnością w GUM. Różnica to dokładnie  $\sqrt{2 \ln 2}$ , czyli około 18%. Nie jest to dużo i w sumie pozwala to na dalsze studia nad niepewnością bez popełniania znaczących błędów. Oczywiście można w tym miejscu wspomnieć jedynie słuszne, zaakceptowane powszechnie przyjęte formuły z GUM, które każdy, kto ma bardziej zaawansowany kalkulator i chęć do rachunków, może zawsze zastosować.

Uczniowie są w stanie określić „połowę szerokości w połowie wysokości” narysowanych przez siebie histogramów i ich najbardziej prawdopodobne wartości odczytane z wykresów. Wypisanie tych wszystkich wyników na tablicy prowadzi do dość zaskakującego, na pierwszy rzut oka, rezultatu. Wszystkie wyznaczone szerokości są bardzo bliskie. Pokazuje to na poprawność zastosowanej metody i pozwala przedyskutować jeszcze raz ewentualne zaobserwowane rozbieżności. Ważne jest stwierdzenie, że pozornie subiektywnie podejście do danych prowadzi do rezultatów całkiem obiektywnych. Wyniki otrzymane na lekcji pokazuje Tabela 1.

Tabela 1. Pomiary swobodnego spadku, wyznaczona z histogramu wartość średnia  $\bar{t}$  i rozrzut pojedynczych pomiarów.

Położenie maksimum	0.79	0.80	0.80	0.80	0.82
	0.78	0.80	0.79	0.80	0.78
	0.80	0.79	0.82	0.80	0.78
	0.78	0.80	0.80	0.81	0.78
Połowa szerokości w połowie wysokości	0.09	0.10	0.12	0.09	0.12
	0.10	0.10	0.10	0.09	0.10
	0.08	0.09	0.07	0.12	0.09
	0.10	0.11	0.09	0.08	0.13



Rysunek 3. Histogram dla wyników swobodnego spadku i trzy linie najlepiej dopasowanych krzywych Gaussa (patrz tekst).

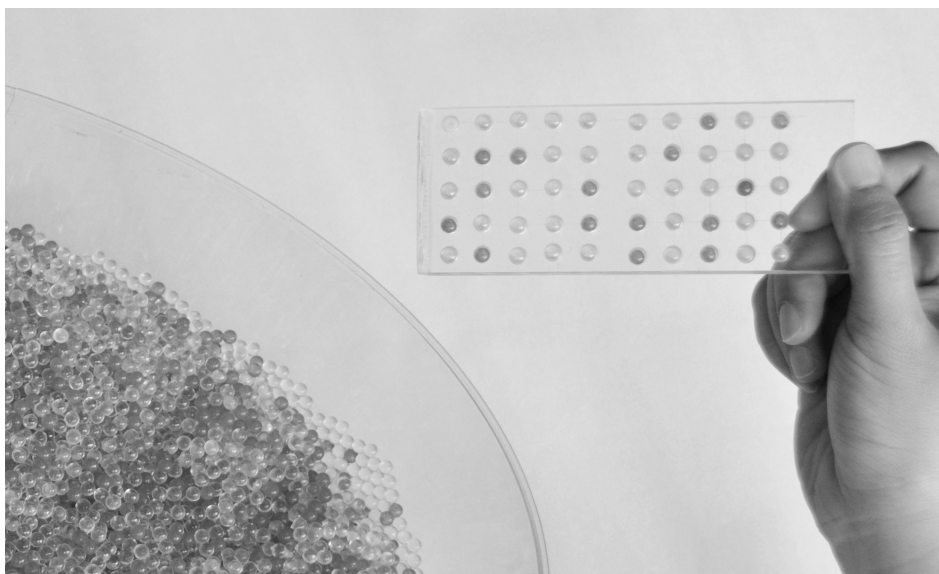
W tym miejscu warto porównać wyniki otrzymane przez dzieci z „profesjonalnymi” dopasowaniami wykonanymi z użyciem oprogramowania PAW (Physics Analysis Workstation, poprzednika pakietu ROOT używanego w CERNie przez naukowców m.in. poszukującymi boskiej cząstki i ciemnej materii). Wyniki tych dopasowań pokazane są na Rys. 3. Trzy linie pokazują „najlepsze dopasowania” do histogramu czasu swobodnego spadku. Pierwsza, pokazana linią kropkowaną to wynik minimalizacji  $\chi^2$  Pearsona, co jest zasadniczo błędne przy tak małej statystyce. Dwie pozostałe linie (ciągła i kreskowana) otrzymano za pomocą metody maksymalizacji wiarygodności, a różnica pomiędzy nimi bierze się z tego, że za pierwszym razem (linia ciągła) uwzględniono wszystkie zmierzone wartości, a za drugim (linia kreskowana) pominięto punkt o wartości czasu większej niż 1.2 s. Różnice w dopasowaniach wydają się istotne. Lekcja, jaką należy stąd wynieść jest jednak taka, że nie ma czegoś takiego, jak uniwersalna recepta na „najlepsze dopasowanie”. Zależy ono w istocie także od „niefizycznych” parametrów, od tego wszystkiego, co wiemy o pomiarze i warunkach w jakich był dokonany. Jest to, nawiasem mówiąc, doskonały przykład ilustrujący sens bayesowskiego sposobu myślenia o prawdopodobieństwie. W naszym, szczególnym przypadku używając standardowych metod klasycznego rachunku prawdopodobieństwo powinniśmy raczej odrzucić ten jeden ekstremalny wynik z serii pomiarowej. Przy bayesowskim ujęciu zostawiamy to do indywidualnej decyzji uczniów, którzy i tak w jakiś swój „pozaświadomy” sposób uwzględnili to już kreśląc swoje dopasowane krzywe. Wychodzi na to, że w gruncie rzeczy dzieci myślą po bayesowsku!

Wracając do lekcji, pora zadać kolejne, pozornie trywialne, acz bardzo istotne w gruncie rzeczy pytanie: czy wykonując więcej i więcej pomiarów (czasu spadania) nasza wiedza o tej wielkości (czasie spadania) zwiększa się, czy nie? Jest to pytanie raczej retoryczne. Im więcej zmierzemy, tym wiedzieć powinniśmy więcej. A więc, jeśli wiedza nasza naprawdę rośnie, pora zadać kolejne, zasadnicze pytanie lekcji o statystyce: jak szybki jest ten wzrost? Jaka jest relacja „błędu średniej” do „średniego błędu” pojedynczego pomiaru (w terminologii Błasiaka, nieprawidłowej zdaniem GUM).

### **Takoyaki**

Aby odpowiedzieć na to pytanie, pozostając konsekwentnym, powinniśmy zadać je Naturze, czyli przeprowadzić eksperyment. Natura, czyli otaczająca nas Rzeczywistość jest złożona i pytanie postawić musimy umiejętnie, aby odpowiedź dała się zauważyć i zanalizować bez stosowania wyrafinowanych metod matematycznych. Aby to osiągnąć musimy Rzeczywistość odpowiednio spreparować, uprościć. Potrzebujemy modelu rzeczywistości, modelu najprostszego, ale jednocześnie na tyle złożonego, by móc studiować na nim statystyczne prawa rządzące procesem pomiaru. Chcielibyśmy badać jedynie niepewności statystyczne (typu A), więc model powinien minimalizować możliwość popełnienia błędów związaną z nieprawidłowym odczytem, niedoskonałością eksperymentatora

(naszych dzieci) i ogólnie wszystkiego, co może podpadać pod niepewności typu B, niestatystyczne.



Rysunek 4. Takoyaki, przyrząd do doświadczeń z kulkami.

Wydaje się, że znaleźliśmy taki model. Co ciekawsze, model ten zna każdy, bo znaleźć możemy go w nieco innej formie w bajce o Kopciuszku. Został on tam stworzony przez złośliwą macochę Kopciuszka, a składa się z miski zmieszanej z popiołem soczewicy. Zakładając w celach statystycznych, że ziarna popiołu są mniej więcej tego samego rozmiaru, co ziarna soczewicy (co nie jest oczywiście prawdą, ale to w końcu bajka tylko) możemy zadać Kopciuszkowi zupełnie poważne, naukowe pytanie: „Jaki jest ułamek ziaren soczewicy, spośród wszystkich ziaren w misce (ziaren popiołu i soczewicy)?”, czyli inaczej: „Jakie jest prawdopodobieństwo tego, że losowo wyciągnięte przez nas ziarenko będzie ziarenkiem soczewicy, a nie popiołu?” Na dobrą sprawę jest to jedyne rozsądne pytanie, jakie możemy zadać Kopciuszkowi, jeśli ograniczyć się do pytań o zawartość miski. Oczywiście każdy może w tym miejscu postąpić metodą Kopciuszka i rozdzielić wszystko, co ma w misce na dwie kupki: soczewicy i popiołu, a potem przeliczyć, ile dokładnie jest jakich ziarenek. W ten sposób oczywiście uzyska się prawdziwą odpowiedź co do zawartości miski, ale jest to sposób dalece „nienaukowy” i niewystarczający, jeśli chodzi o pytanie natury ogólnej. Fizycy badając Rzeczywistość tak nie postępują. Nie da się bowiem wykonać wszystkich możliwych doświadczeń i poznać zawczasu odpowiedź na każde pytanie. Nie ma nawet sensu próbować! Możliwe jest natomiast pobranie próbki Rzeczywistości, próbki niewielkiej, ale reprezentatywnej. Można tę próbkę przeliczyć (bo jest właśnie niewielka) i na podstawie tego wyniku

można wnioskować w zawartości całej miski Kopciuszka – naszego modelu Rzeczywistości.

W naszym doświadczeniu moglibyśmy użyć oczywiście także popiołu i soczewicy, ale choćby ze względu na to, że trudno jest o popiół nie pozostawiający po lekcji śladów na eksperymentatorach, postanowiliśmy użyć małych, szklanych kulek. Ponieważ musimy mieć w modelu co najmniej dwa rodzaje kulek, część z nich – około 20% – pomalowaliśmy sprayem na jasnozielony kolor i w ten sposób stworzyliśmy model nasz model Rzeczywistości – Rzeczywistość Pokolorowaną. Jedyne pytanie jej dotyczące, jakie możemy zadać to: „Jaki ułamek kulek został pomalowany?” I tym właśnie się zajmijmy w sposób ścisły i naukowy.

Na lekcji każda grupa dzieci (zwykle około pięciu) wyposażona zostaje w próbkę Rzeczywistości do badań (pojemnik zawierający kilka tysięcy kulek wybranych losowo z większego pojemnika – Rzeczywistości Pokolorowanej). Pierwsze zadanie, jakie przed nimi stawiamy to spróbować ustalić możliwie dokładnie, jaki jest ułamek zielonych kulek, jeśli wnioskowanie ograniczyć do przeliczania zacerpnętej małej próbki Rzeczywistości Pokolorowanej, powiedzmy próbki o liczebności 25 kulek.

Aby eksperyment przebiegał sprawnie i precyzyjnie wynaleźliśmy specjalny przyrząd do pobierania próbki 25 (50) kulek. Jest to fragment grubej płytki z plexiglasu, w której nawierciliśmy 25 (naprawdę 2\*25) wgłębień o rozmiarze nieco większym niż średnica kulek i takiej samej głębokości. Od nazwy japońskich ciasteczek przygotowywanych na podobnych, lecz oczywiście większych, blachach nazwaliśmy nasz przyrząd „takoyaki pan”, w skrócie: „takoyaki”. Pokazuje go Rys.4. Po zanurzeniu takoyaki w Rzeczywistości, kulki wpadają same we wgłębienia i łatwo można je policzyć. Pojedynczy pomiar polega na podaniu liczby zielonych kulek w 25 wgłębieniach takoyaki.

W pierwszym eksperymencie z użyciem tokayaki prosimy uczniów o wykonanie 10 takich pomiarów. Służy on także w opanowaniu techniki mierzenia i oswojeniu się z „przyrządem pomiarowym”. Wszystkie wyniki wszystkich grup zapisujemy na tablicy, co w sumie daje serię około 50 wyników, wystarczająco wiele, aby sporządzić histogram. Uczniowie mają go teraz samodzielnie w grupach narysować, dopasować doń najlepszą ich zdaniem krzywą Gaussa i oszacować jej parametry. Ponieważ jest to powtórzenie procedury już raz zastosowanej, opracowanie tych danych przebiega na ogół sprawnie i asysta potrzebna jest już tylko w nielicznych przypadkach. Po niedługim czasie możemy porównać wyniki uzyskane przez grupy, tak dotyczące położenia maksimum, jak i **połowy szerokości w połowie wysokości**. Należy zwrócić uwagę na ewentualne wyniki odbiegające od innych. Dobrze jest przypadki te przedyskutować i wyjaśnić ewentualne rozbieżności tak, by wszyscy zauważyli niuanse i możliwe pułapki procedury dopasowania krzywej.

W tym miejscu można, choć nie jest to konieczne i zależy to od ich stopnia zainteresowania i wiedzy ogólnomatematycznej, wspomnieć o różnicy pomiędzy

naszą miarą rozrzutu (**połowa szerokości w połowie wysokości**) a odchyleniem standardowym znanym z niektórych podręczników (i GUM!). Po pierwsze, rozkład liczby zielonych kulek nie jest normalny. Jest to rozkład dyskretny i oczywiście nigdy nie osiąga wartości mniejszych od 0. Dość dobrym przybliżeniem jest tu rozkład Poissona, choć i to nie jest do końca prawdą. Oczywiście jest, że liczba zielonych kulek nigdy nie przekracza liczebności całej próbki (25 w tym doświadczeniu), którego to ograniczenia w rozkładzie Poissona nie ma, ciągnie się on do nieskończoności. Rozkład w naszym doświadczeniu jest tak naprawdę rozkładem dwumianowym. Mamy do czynienia z ciągiem 25 prób Bernoulliego i interesuje nas całkowita liczba sukcesów (wylosowania zielonej kulki). Dla ułamka zielonych kulek wynoszącego około  $1/5$ , jak w naszym przykładzie, różnica pomiędzy **połową szerokości w połowie wysokości** w rozkładzie Gaussa i dwumianowym wynosi około 11%.

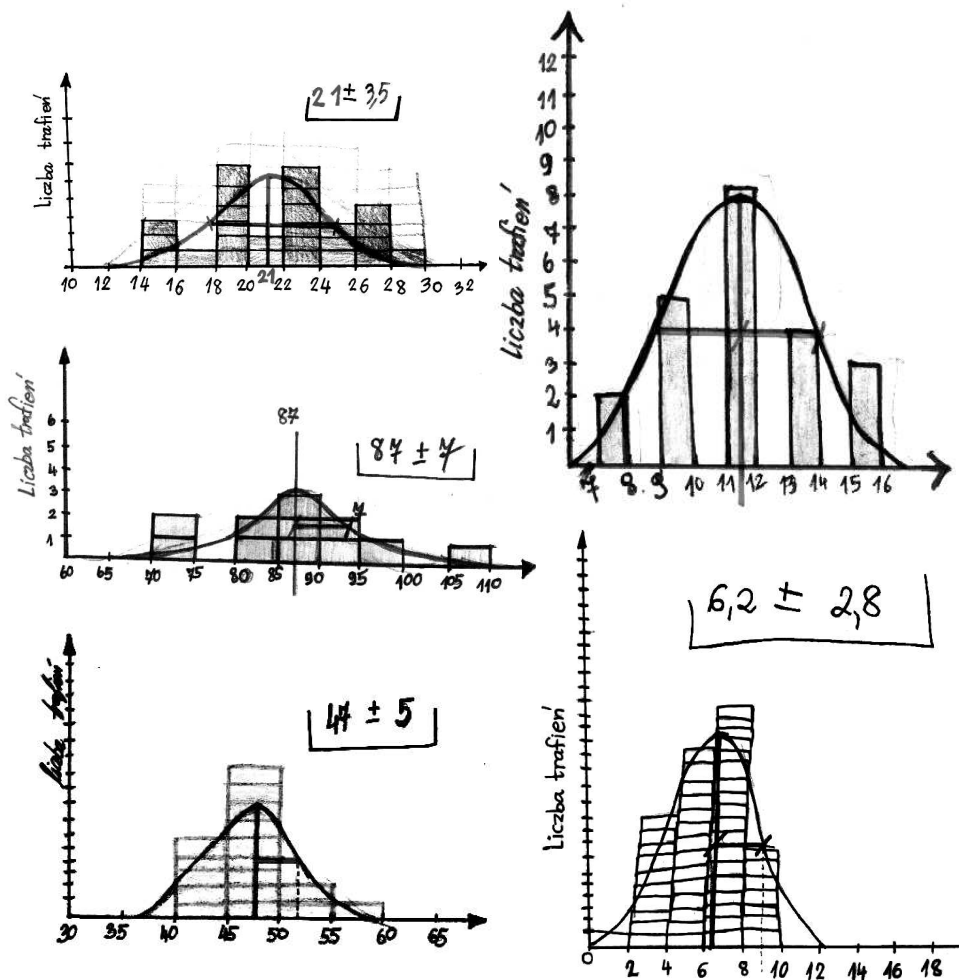
Niezależnie od tej dygresji zapamiętajmy wartości „położeń maksimum” i „**połówek szerokości w połowie wysokości**” do użycia w późniejszej, globalnej analizie, stawiamy uczniom zasadnicze pytanie: „Jaka będzie szerokość histogramów, gdy użyjemy większych próbek?”

Ustalmy, że wykonamy odpowiednie pomiary dla liczebności 50, 100 200 i 400 kulek w próbie. Wynik dla 400 kulek uzyskujemy powtarzając ośmiokrotnie pomiary 50 ( $2 \cdot 25$ ) kulek na takoyaki i sumując liczby kulek zielonych. Jest to dość żmudne zajęcie, dlatego warto, aby ten pomiar dokonywały dwie grupy uczniów i budowały jeden wspólny histogram. Statystykę jaką zebrać mają uczniowie ustalamy biorąc pod uwagę sprawność i czas jaki mamy do dyspozycji. Pomiar dla liczebności 400 jest bardzo ważny i może się okazać krytyczny dla dalszej analizy, dlatego należy dołożyć wszelkich starań, aby zebrać statystykę co najmniej kilkunastu przypadków.

Wyniki otrzymane na jednej lekcji pokazuje Rys. 5, a wartości liczbowe przedstawia Tabela II.

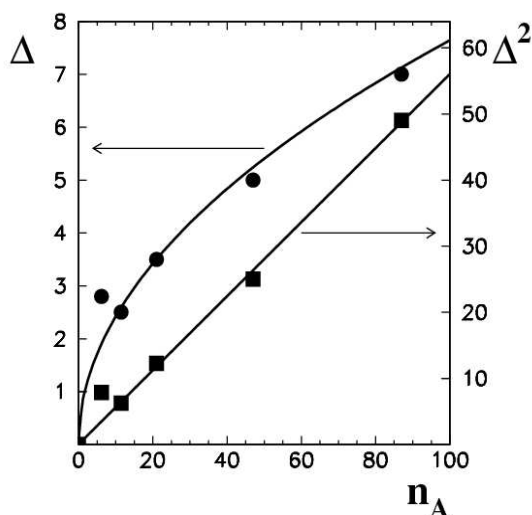
Tabela II. Położenie maksimum  $n_A$  i rozrzut wyników pojedynczych pomiarów dla rozkładów otrzymanych przy różnych wielkościach próbek. Drugi rząd pokazuje liczbę pomiarów wykonaną dla danej statystyki

wielość próbki ( $n$ )	25	50	100	200	400
statystyka ( $N$ )	50	22	27	17	11
położenie maksimum ( $n_A$ )	6.2	11.5	21.0	47	87
rozrzut ( $\Delta$ )	2.8	2.5	3.5	5	7
$f = n_A / n$	0.248	0.23	0.21	0.235	0.218



Rysunek 5. Wyniki dopasowywania krzywych Gaussa do histogramów otrzymanych dla różnej liczebności próbek.

Ostatni etap działalności doświadczalnej obejmuje sporządzenie jednego wykresu, na którym umieścimy wszystkie uzyskane dotąd wyniki w postaci „połowy szerokości w połowie wysokości” (oznaczyliśmy to jako  $\Delta$ ) w funkcji „położenia maksimum” ( $n_A$ ) znalezionych dla próbek 25, 50, 100, 200 i 400 kulek. Pokazany jest on na Rys.6. (kółka i skala na lewej osi).



Rysunek 6. Połowa szerokości w połowie wysokości (kółka i skala lewa) jako funkcja położenia maksimum, oraz jej kwadrat (kwadraty i skala prawa). Linie pokazują gładką spodziewaną (kwadratową i liniową) zależność dopasowaną do danych.

Dodatkowo na wykresie umieszczamy punkt  $(0, 0)$ , ponieważ wiadomo jest, że w próbie zawierającej zero zielonych kulek zawsze otrzymamy ten sam wynik, a więc odpowiednia krzywa Gaussa będzie miała „zerową szerokość” – żadnej niepewności co do wyniku! Każda teoretyczna zależność  $\Delta$  od  $n_A$  musi przechodzić przez punkt  $(0, 0)$ .

Jeśli zapytamy teraz uczniów o to, jaka jest zależność  $\Delta$  od  $n_A$ , nikt, kto nie zna właściwej odpowiedzi z innych, wcześniejszych studiów i źródeł podręcznikowych, nie jest w stanie wpaść na nią bez dodatkowej podpowiedzi. Można próbować nakreślić jakąś gładką krzywą, jednak odgadnięcie zależności funkcyjnej z Rys. 6 jest na pierwszy rzut oka praktycznie niemożliwe. Jedyne, co daje się zauważyć, to to, że nie jest to zależność liniowa. Nie da się poprowadzić prostej zgodnej z obserwowaną w danych tendencją – zależność „zagina się”. Podpowiedź ze strony nauczyciela jest całkowicie w tym miejscu uzasadniona (nauczyciel „wie lepiej”!), oszczędza ona po prostu cenny czas.

Jeśli nie liniowa, to może kwadratowa?

Uczniowie znają parabolę w postaci „ $y = x^2$ ”, czyli funkcję, która rośnie szybciej i szybciej (ruch jednostajnie przyspieszony!) Obrócenie tablicy o 90 stopni zapewne by pomogło, ale zamiast tego najprościej będzie sporządzić wykres zależności  $\Delta^2$  od  $n_A$ . Można to zrobić na tym samym wykresie przeskalowując

odpowiednio oś pionową, tak jak to pokazuje skala po prawej stronie rysunku. Odpowiadające jej punkty  $(\Delta^2, n_A)$  zaznaczona kwadratami. Jak widać przyjęcie, że  $\Delta^2$  zależy linowo od  $n_A$  jest całkiem dobrze ugruntowane eksperymentalnie. To właśnie stwierdzenie jest zasadniczym wynikiem wszystkich działań pomiarowych na naszym modelu rzeczywistości. Pozornie nie widać w nim nic nadzwyczajnego, jednak warto poświęcić chwilę na prześledzenie niezwykle istotnych, ukrytych w nim konsekwencji.

### Konsekwencje

Znaleziona proporcjonalność  $\Delta^2$  do  $n_A^1$  pozwala nam na uczynienie istotnego kroku w obliczeniu „błędu średniej”.

Średnia (wartość spodziewana, czy najbardziej prawdopodobna) liczba zielonych kulek w próbie 25 kulek na takoyaki nie jest w zasadzie tym, o co chcieliśmy początkowo zapytać Naturę. Chodziło nam o ułamek zielonych kulek w ogóle, czyli  $f = n_A / n$ , gdzie  $n$  oznacza wielkość próbki. Ułamek taki oczywiście fluktuuje, zmienia się, raz jest taki, a przy powtórzeniu wykonaniu eksperymentu inny i jego rozrzut od doświadczenia do doświadczenia (przy stałym  $n$ ) jest określony przez zmierzony rozrzut liczby kulek zielonych

$$\Delta = \frac{\Delta_A}{n} = \frac{\sqrt{n_A}}{n} = \frac{\sqrt{n_A}}{n_A/f} = \frac{f}{\sqrt{n_A}} \quad (1)$$

W tabeli II mamy w piątym rzędzie podane pięć wyznaczonych wartości ułamka  $f$  otrzymanych przez różne grupy uczniów dla różnych wielkości próbek. Oznaczmy je  $f_j$

$$f_j = \frac{\sum_i n_{Aij}}{\sum_i n_{ij}} = \frac{n_{Aj}}{n_j} \quad (2)$$

gdzie  $n_{Aij}$  to liczby zielonych kulek w grupie  $j$ -tej w doświadczeniu o numerze  $i$ -tym. Dla porządku przyjęliśmy definicję całkowitej liczby zielonych kulek w grupie:  $n_{Aj} (\sum_i n_{Aij})$  i całkowitej liczby wszystkich kulek w grupie:  $n_j = \sum_i n_{ij}$ .

---

<sup>1</sup> Proporcjonalność, nie równość, której można by się spodziewać w tym miejscu, bierze się ze wspomnianych poczynionych przez nas uproszczeń. Zastąpienie sigmy w rozkładzie normalnym **połową szerokości w połowie wysokości** i rozkładu dwumianowego rozkładem normalnym dają odpowiednio zaniżenie miary rozrzutu o 18% i 11%. Prowadzi to do współczynnika proporcjonalności równego  $[(1-0.18) * (1-0.11)]^2 \approx 0.53$ , co zgodne jest z wynikami przedstawionymi na Rys. 6.



Teraz pora na zastanowienie się, jak wykorzystać **wszystkie** zebrane w doświadczeniach wszystkich grup informacje o  $f_j$  do najlepszego określenia globalnej wartości ułamka  $f$ . Praktycznie wszyscy uczniowie biorący udział w lekcjach, podobnie jak i najpewniej zdecydowana większość ludzi zsumowałyby pięć liczb z ostatniego wiersza Tabeli II i wynik podzieliła przez 5, bo tyle tych liczb tam jest. W wyniku otrzymalibyśmy wartość globalnego  $f$  równą „średniej” 0.2282 i oczywiście nie jest to poprawna odpowiedź na postawione wyżej ogólne pytanie. Dla najbardziej precyzyjnego określenia wartości  $f$  powinniśmy przecież niewątpliwie dodać do siebie ilości zielonych kulek zaobserwowane za każdym razem w każdej grupie i w każdym doświadczeniu i wynik takiego sumowania podzielić przez całkowitą liczbę wszystkich kulek jakie użyliśmy we wszystkich doświadczeniach

$$\langle f \rangle = \frac{\sum_k n_{Ak}}{\sum_j n_j} \quad (3)$$

(różnica pomiędzy  $k$  i  $j$  jest żadna, są to indeksy sumowania i jak by je nie nazwać, chodzi nam zawsze o to samo – sumowanie po wszystkich grupach). Wielkości  $n_{Ak}$  w równaniu (3) można zastąpić na mocy równania (2) odpowiednim iloczynem ( $n_{Ak} = f_k * n_k$ ), a następnie wprowadzić możemy wielkości, które nazwiemy **wagami**  $w_j$ .

$$\langle f \rangle = \frac{\sum_k n_k}{\sum_j n_j} = \sum_k f_k \frac{n_k}{\sum_j n_j} = \sum_{k=1}^N f_k w_k \quad (4)$$

Wagi  $w_k$  są dla naszych rozważań niezwykle istotne. Określają one jak ważny (stąd nazwa) jest dany wynik  $f_k$ . Jak widać z równania (4) są równe w naszym doświadczeniu procentowemu udziałowi liczby kulek analizowanej w danej grupie. Im więcej kulek przetestujemy, tym nasz wynik jest ważniejszy. Jest to konkretna i precyzyjna odpowiedź na nasze pytanie. Niczego więcej nam nie potrzeba.

W przypadku identycznych eksperymentów, gdy wszystkie wyniki są jednakowo dokładne, wszystkie wagi są takie same ( $n_j = n_0 = \text{const}$ ),

$$w_j = \frac{n_0}{\sum_{j=1}^N n_0} = \frac{1}{N} \quad (5)$$

i wtedy

$$\langle f \rangle = \sum_{j=1}^N f_j w_j = \frac{f_1 + f_2 + \dots + f_N}{N} \quad (6)$$

i otrzymujemy w wyniku **średnią arytmetyczną**. Jak z tego widać średnia arytmetyczna jest dobra jedynie wtedy, gdy składamy wyniki eksperymentów, które mają tę samą niepewność. Jeśli jest tak, jak w przypadku naszego eksperymentu z kulkami, gdzie niepewności są różne, jak to pokazuje w naszym przypadku Tabela III, musimy posługiwać się wzorami dokładnymi (4). Wtedy wartość ułamka  $f$  rzeczywiście najlepiej zmierzona przez nas wynosi 0.225. Różnica nie jest duża, bo wynosi około 1%, W przypadkach bardziej skomplikowanych, gdzie różnice niepewności są naprawdę duże nieprawidłowe użycie średniej arytmetycznej może prowadzić do znaczących błędów (błędów, nie niepewności!) w analizie i w konsekwencji do wyciągnięcia błędnych wniosków.

Tabela 3. Wagi wyników każdej grupy  $\sum_j \sum_i n_{ij} = 12850$ . I inne wielkości pomocne w obliczeniach ostatecznych.

liczebność próbki $n$	25	50	100	200	400
statystyka $N$	50	22	27	17	11
$N * n$	1250	1100	2700	3400	4400
waga $w_j$	0.10	0.09	0.21	0.265	0.342
$f = nA/n$	0.248	0.23	0.21	0.235	0.218
				$\langle f \rangle = 0.225$	
$\Delta_i$	0.0063	0.0068	0.0043	0.0039	0.0034
				$\Delta = 0.002$	

### Uogólnienie

Wagi zdefiniowane w równaniu (4) można przepisać zastępując wartości  $n$  ilorazami  $n_A$  i ułamka  $f$  na mocy równania (2)

$$w_j = \frac{n_j}{\sum_k n_k} = \frac{n_{Aj} / f_j}{\sum_k n_k} \quad (7)$$

Wartości  $f_j$  ( $f_k$ ) określone przez każdą z grup są w zasadzie bardzo podobne i przy szacowaniu wag możemy uznać, że są niemal identyczne. Jeśli zastąpić je stałą (na przykład  $\langle f \rangle$ ) skrócą się one i w liczniku i mianowniku pozostaną jedynie odpowiednie wartości  $n_A$ .

W tym momencie możemy zastosować odkrytą przez nas doświadczalnie zależność proporcjonalności  $n_A$  do  $\Delta$  i po prostych przekształceniach otrzymujemy:

$$w_j = \frac{n_{Aj}}{\sum_k n_{Ak}} = \frac{1/\Delta_j^2}{\sum_k 1/\Delta_k^2} = \frac{1}{\Delta_j^2} / \left( \sum_k 1/\Delta_k^2 \right) \quad (8)$$

Suma w mianowniku jest taka sama dla wszystkich grup i jest jedynie czynnikiem normalizacyjnym, zapewniającym, że suma wszystkich wag jest jednością – tak musi być z zasady zachowania prawdopodobieństwa. Najistotniejsze we wzorze (8) jest to, że wagi są odwrotnie proporcjonalne do kwadratów niepewności pomiarów do jakich się odnoszą. Warto zaznaczyć po raz kolejny, że ten bardzo podstawowy wynik w istocie został uzyskany przez uczniów na drodze doświadczalnej. W naszym eksperymencie tokayaki sami oni odkryli jedno z podstawowych praw Natury. Przy okazji zauważmy, że rozwiązanie to jest jak najbardziej zgodne z ideami zawartymi w GUM.

### Błąd średniej

Niepewność średniej arytmetycznej może być wyprowadzona bez specjalnych trudności na drodze elementarnej w rozumowaniu podobnym do przedstawionego powyżej. Z równania (1) otrzymujemy:

$$\frac{f_j^2}{\Delta_j^2} = n_{Aj} \quad (9)$$

Niepewność wyniku końcowego  $f$  spełnia także to samo równanie:

$$\frac{f^2}{\Delta^2} = \sum_j n_{Aj} \quad (10)$$

Stosując przybliżenie  $f_j \approx f$  i porównując wzory (9) i (12) otrzymujemy:

$$\frac{1}{\Delta^2} = \sum_j \frac{1}{\Delta_j^2} \quad (11)$$

Możemy użyć tego wyniku do danych zawartych w Tabeli III i oszacować niepewność naszej średniej (ważonej) wartości  $f$ . W sumie analiza wszystkich naszych eksperymentów takoyaki prowadzi do rezultatu:  $f = 0.255 \pm 0.02$ .

Tym samym znaleźliśmy odpowiedź na postawione na początku lekcji pytanie. Potrafimy wyznaczyć poprawnie najlepszą wartość ułamka  $f$  i jesteśmy w stanie podać niepewność tego wyniku.

Wracając do prostego przypadku składania wyników eksperymentów jednakowo dokładnych z wzoru (11) otrzymujemy wynik:

$$\Delta = \frac{\Delta_s}{\sqrt{N}} \quad (12)$$

gdzie  $\Delta_s$  oznacza niepewność wyniku pojedynczego pomiaru, a  $\Delta$  jest niepewnością średniej z  $N$  powtórzeń tego eksperymentu.

Tego wyniku możemy teraz z kolei skorzystać w przypadku dowolnego innego, powtarzanego wiele razy, eksperymentu polegającego niekoniecznie na zliczaniu kulek. Na początku lekcji przeprowadziliśmy pomiary czasu spadania i możemy je teraz zanalizować spójnie i kompleksowo. Najlepsze oszacowanie tego czasu to  $(0.80 \pm 0.016)$  s. Po przełożeniu tego na wartość przyspieszenia ziemskiego (pomijamy drobne niuansy związane z nieliniowością rozpatrywanej funkcji zmiennej losowej, co przy niewielkich odchyleniach jest uzasadnione) otrzymujemy  $(9.4 \pm 0.2)$  m/s<sup>2</sup>. Zupełnie niezły wynik!

### Podsumowanie

Pokazaliśmy, że prawidłowe (GUM) podejście do statystycznej analizy danych może być z powodzeniem przekazane uczniom w wieku licealnym i gimnazjalnym zaczynającymi dopiero przygodę z empirycznym poznawaniem praw Natury. Budowanie histogramów i użycie ich potem do określania niepewności pomiarowej jest na tym etapie rozwoju całkiem naturalne. Wyprowadzenie zależności pomiędzy „średnim błędem” pojedynczego pomiaru a „błędem średniej” z wielu pomiarów (wzór 12) posiłkując się doświadczeniem daje solidne podstawy do poprawnego zrozumienia, a nie tylko nauczenia się, tej i innych procedur niezbędnych w naukach przyrodniczych.

### Bibliografia

- International Organization for Standardization*, Geneva, Guide to the Expression of Uncertainty in Measurement, (1995).
- A. Buffler and S. Allie, *Teaching Measurement and Uncertainty the GUM Way*, The Physics Teacher, 46, 539{543, (2008).
- General Requirements for the Competence of Calibration and Testing Laboratories ISO*, ISO/IEC 17025:1999, Geneva (1999).
- Guide To The Expression Of Uncertainty In Measurement*, ISO, Geneva, Switzerland (1993).
- The Expression of Uncertainty and Confidence in Measurement*, United Kingdom Accreditation Service, M3003 Edition 2, January (2007).
- Quantifying Uncertainty in Analytical Measurement*, EURACHEM/CITAC Guide CG 4, Eds. S. L. R. Ellison, M. Rosslein, A. Williams, 2nd edition, (2000).
- B.N. Taylor and C.E. Kuyatt, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, National Institute of Standards and Technology, United States Department of Commerce Technology Administration, Technical Note 1297 (1994).
- R. Lippmann-Kung, *Teaching the Concepts of Measurement: One example of a concept-based laboratory course*, American Journal of Physics, 73, 771{777 (2005).
- W. Błasiak, *Error and Uncertainty in Physical measurement*, Physics Education, 18, 290, (1993).

- G. Alto and F. Agnoli, *Little statisticians grow up: Exploring the concept of variability in young children*, The International Conference on Teaching Statistics-8, Ljubljana, Slovenia, (2010).
- S. Kazak and J. Confrey, *Elementary School Students' Informal and Intuitive Conception of Probability and Distribution*, The International Conference on Teaching Statistics-7, Salvador, Brazil, (2006).
- R. Delmas and Y. Liu, *Exploring Students' Conception of the Standard Deviation*, Stat. Edu. Research Journal 4, 55{82, (2005).
- J.O. Moya and A.E. Ortega, *Meaning of the Dispersion and its Measures in Secondary Education*, The International Conference on Teaching Statistics-7, (2006).
- R.P. Feynman, R.B. Leighton, M. Sands, *Feynman lectures in Physics*, 1, New York, Addison-Wesley, (1963).
- K. Pearson, *The problem of the Random Walk*, Nature, 72, 294, July 27, (1905).
- Lord Rayleigh, *The problem of the Random Walk*, Nature, 72, 318, August 3, (1905).
- K. Pearson, *The problem of the Random Walk*, Nature, 72, 342, August 10, (1905).
- I. Duerdoth, *Teaching uncertainties*, Phys. Educ. 44, 138, (2009).

