

Analiza niepewności pomiarowych

Wstęp do analizy danych

Wykład 4

Bezpośredni pomiar pojedynczej wielkości fizycznej

1 Uwagi wstępne

Podczas poprzednich wykładów zostały przypomniane podstawowe pojęcia rachunku prawdopodobieństwa, którym posługujemy się do opisu błędów przypadkowych. Traktujemy rachunek prawdopodobieństwa jak sprawne narzędzie przygotowane przez matematyków. Nasze potrzeby ograniczą się do obliczania wartości oczekiwanych zmiennych losowych oraz prawdopodobieństw przyjmowania przez badane zmienne losowe wartości mieszczących się w zadanych przedziałach. W rozważaniach pojawiają się tylko zmienne dyskretne (tj. o przeliczalnym zbiorze wartości) oraz zmienne ciągłe, ale tylko takie, których gęstości prawdopodobieństwa zadane są funkcjami ciągłymi. Przykładami dyskretnych zmiennych losowych są np. liczba sukcesów w ciągu n prób Bernoulliego, czy pierwiastek z liczby oczek na górnej ścianie kostki do gry po kolejnym rzucie. Z przykładem zmiennej losowej ciągłej spotykamy się w naszym modelu błędu przypadkowego, gdzie zakładamy, że zmienna losowa *błąd przypadkowy* może przyjmować dowolne wartości rzeczywiste i podlega rozkładowi Gaussa (patrz niżej). Rozkład Gaussa jest najczęściej przyjmowanym modelem rozkładu błędu przypadkowego i zawsze, gdy nie ma przesłanek, żeby posługiwać się innym rozkładem, stosowany jest rozkład Gaussa. Rozkład ten przyjmowany jest także często jako model rozkładu zmienności cech – np. wzrostu członków wybranej populacji (o ile jest wystarczająco liczna). Gęstość rozkładu Gaussa jest funkcją ciągłą, szybko malejącą do zera, gdy jej argument dąży do $\pm\infty$. Funkcji pierwotnej (całki nieoznaczonej) gęstości rozkładu Gaussa nie można zapisać w postaci wzoru analitycznego zawierającego funkcje elementarne, ale arkusze kalkulacyjne i pakiety programów do obliczeń numerycznych zawierają algorytmy pozwalające obliczać wartości całek z gęstości rozkładu Gaussa w dowolnych granicach. Rozkład Gaussa definiują dwa parametry: wartość oczekiwana μ i dyspersja σ . Wartość oczekiwaną utożsamiamy z dokładną wartością mierzonej wielkości, a dyspersję z dokładnością, z jaką tę wielkość jesteśmy w stanie zmierzyć, czyli z niepewnością pomiaru.

2 Model rozkładu prawdopodobieństwa wartości błędów przypadkowych

W §11 *Slajdów/transparencji do wykładu* sformułowany został model rozkładu prawdopodobieństwa błędów przypadkowych jako granicy rozkładu prawdopodobieństwa dla n czynników losowych,

z których każdy powoduje odchylenie wyniku pomiaru od wartości dokładnej o $+\varepsilon$ lub $-\varepsilon$. Obie wartości przyjmowane są z jednakowym prawdopodobieństwem ($p = q = \frac{1}{2}$). Dokonaliśmy przejścia do granicy, $n \rightarrow \infty$, $\varepsilon \rightarrow 0$ z dodatkowym warunkiem $n\varepsilon^2 = \sigma^2$ (σ^2 jest wielkością stałą) i otrzymaliśmy gęstość rozkładu prawdopodobieństwa $f(x; \mu, \sigma)$ uzyskania jako wynik pomiaru wartości x , gdy dokładna wartość wynosi μ :

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right).$$

Funkcja $f(x; \mu, \sigma)$ nazywana jest funkcją Gaussa, a definiowany za jej pomocą rozkład nazywany jest **rozkładem Gaussa**. Wartość czekiwana zmiennej x podlegającej rozkładowi Gaussa równa się μ :

$$\mathcal{E}(x) = \int_{-\infty}^{\infty} x f(x; \mu, \sigma) dx = \int_{-\infty}^{\infty} (x - \mu) f(x; \mu, \sigma) dx + \int_{-\infty}^{\infty} \mu f(x; \mu, \sigma) dx = \mu,$$

a wariancja wynosi σ^2 :

$$\mathcal{E}((x - \mu)^2) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x; \mu, \sigma) dx = \sigma^2.$$

Model Gaussa jest powszechnie przyjmowanym modelem dla rozkładu błędów przypadkowych. Będziemy się do niego wielokrotnie odwoływać. W wielu przypadkach wystarczy jednak założenie, że rozkład prawdopodobieństwa błędu przypadkowego ma dobrze określone i skończone: wartość oczekiwaną i wariancję. Takie założenie wystarczy nam do zbudowania wyrażeń pozwalających wyznaczać oceny wartości oczekiwanej i wariancji rozkładu na podstawie skończonej, losowej próby – serii niezależnych, równoważnych, pomiarów.

3 Bezpośredni pomiar pojedynczej wielkości fizycznej

Jeśli wyeliminowane zostały błędy grube i systematyczne, to wartość oczekiwaną rozkładu błędów przypadkowych utożsamiamy z wartością dokładną μ mierzonej wielkości, a pierwiastek ze średniego odchylenia kwadratowego, σ , błędów przypadkowych od ich wartości oczekiwanej przyjmujemy za miarę dokładności pomiaru – **niepewność pomiaru**. Wartości μ i σ jednak nie znamy, ale możemy ocenić ile wynoszą na podstawie wyników **serii niezależnych, równoważnych pomiarów**, gdy pomiary wykonywano w tych samych warunkach za pomocą tych samych przyrządów pomiarowych.

Zakładamy:

- znamy wyniki serii n pomiarów o wynikach: x_i , $i = 1, \dots, n$;
- o każdym z pomiarów zakładamy, że: $\mathcal{E}(x_i) = \mu$, $\mathcal{E}((x_i - \mu)^2) = \sigma^2$;
- pomiary są niezależne, skąd wynika, że $\mathcal{E}((x_i - \mu)(x_j - \mu)) = \delta_{ij}\sigma^2$, gdzie δ_{ij} oznacza tzw. deltę Kroneckera: $\delta_{ij} = 1$, gdy $i = j$ i $\delta_{ij} = 0$, gdy $i \neq j$.

Naturalnym kandydatem na wyrażenie pozwalające ocenić (znaleźć możliwie najlepsze przybliżenie) wartości μ jest średnia, $\bar{x}$, ze wszystkich wyników pomiarów:

$$\bar{x} := \frac{1}{n} \sum_{i=1}^n x_i$$

Obliczmy wartość oczekiwaną tego wyrażenia:

$$\mathcal{E}(\bar{x}) = \mathcal{E}\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} \sum_{i=1}^n \mathcal{E}(x_i) = \frac{1}{n} \cdot n\mu = \mu.$$

Skorzystaliśmy z liniowości operacji obliczania wartości oczekiwanej. Sprawdźmy, jak dobrze wielkość $\bar{x}$ przybliża μ – obliczmy średnie odchylenie kwadratowe $\bar{x}$ od μ :

$$\begin{aligned} \mathcal{E}((\bar{x} - \mu)^2) &= \mathcal{E}\left(\left(\frac{1}{n} \sum_{i=1}^n x_i - \mu\right)^2\right) \\ &= \mathcal{E}\left(\left(\frac{1}{n} \sum_{i=1}^n (x_i - \mu)\right)^2\right) \\ &= \frac{1}{n^2} \mathcal{E}\left(\sum_{i=1}^n (x_i - \mu) \sum_{j=1}^n (x_j - \mu)\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \mathcal{E}((x_i - \mu)(x_j - \mu)) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \delta_{ij} \sigma^2 = \frac{\sigma^2}{n} \end{aligned}$$

Znowu skorzystaliśmy z liniowości obliczania wartości oczekiwanej, a równość w drugiej linii otrzymaliśmy podstawiając $\mu = \frac{1}{n} \sum_{i=1}^n \mu$.

Okazało się, że średnia $\bar{x}$ lepiej przybliża μ niż pojedynczy pomiar – jej średnie odchylenie kwadratowe od μ jest n razy mniejsze niż każdego z pomiarów uwzględnionych podczas obliczania tej średniej. Wiemy więc już jak najlepiej ocenić wynik pomiaru. Nie wiemy jednak ile wynosi σ . Jeśli $\bar{x}$ dobrze przybliża μ , to możemy oczekiwać, że suma kwadratów $(x_i - \bar{x})^2$ będzie miała związek z σ^2 . Sprawdźmy to – obliczmy wartość oczekiwaną tej sumy:

$$\begin{aligned} \mathcal{E}\left(\sum_{i=1}^n (x_i - \bar{x})^2\right) &= \mathcal{E}\left(\sum_{i=1}^n (x_i - \mu + \mu - \bar{x})^2\right) \\ &= \mathcal{E}\left(\sum_{i=1}^n ((x_i - \mu)^2 + (\bar{x} - \mu)^2 - 2(x_i - \mu)(\bar{x} - \mu))\right) \\ &= \sum_{i=1}^n \mathcal{E}((x_i - \mu)^2) + \sum_{i=1}^n \mathcal{E}((\bar{x} - \mu)^2) - \frac{2}{n} \sum_{ij=1}^n \mathcal{E}((x_i - \mu)(x_j - \mu)) \\ &= (n-1)\sigma^2 \end{aligned}$$

Mamy więc wyrażenie pozwalające ocenić wartość σ^2 na podstawie wyników serii niezależnych, równoważnych pomiarów:

$$s^2 := \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \Rightarrow \mathcal{E}(s^2) = \sigma^2$$

Dla wariancji (średniego odchylenia kwadratowego) średniej $\bar{x}$ otrzymujemy:

$$s_{\bar{x}}^2 := \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2 \Rightarrow \text{Var}(\bar{x}) = \mathcal{E}(s_{\bar{x}}^2) = \frac{\sigma^2}{n}.$$

Tym samym mamy już najlepsze oceny wartości mierzonej wielkości $\hat{\mu} := \bar{x}$ i niepewności tak uzyskanego wyniku $u(\bar{x}) = s_{\bar{x}} := \sqrt{s_{\bar{x}}^2}$.

Fizycy przeważnie niepewność u pomiaru utożsamiają z pierwiastkiem z wariancji wyniku – jest to tzw. **niepewność standardowa**. W innych dziedzinach często podawana bywa tzw. **niepewność rozszerzona** równa $k \cdot u$ z najczęściej przyjmowaną wartością $k = 2$. Z tego powodu zawsze należy podać jaki rodzaj niepewności podajemy po znaku $\pm$ w zapisie wyniku w postaci $\bar{x} \pm u(\bar{x})$. Więcej szczegółów dotyczących zapisu wyniku zawiera §14 *Slajdów/transparencji do wykładu* i kończąca go Uwaga 1.

Wielkość s^2 (i także $s_{\bar{x}}^2$) jest funkcją zmiennych losowych – wyników pomiarów i w związku z tym sama też jest zmienną losową. Możemy zapytać ile wynosi jej wariancja.

§11 *Slajdów/transparencji do wykładu* kończy wzór podający wartość wariancji zmiennej s^2 (jego wyprowadzenie jest, co do zasady, podobne jak wzoru dla s^2 tylko o wiele bardziej żmudne):

$$\text{Var}(s^2) := \mathcal{E}((s^2 - \sigma^2)^2) = \frac{1}{n} \mathcal{E}((x_i - \mu)^4) - \frac{n-3}{n(n-1)} \sigma^4.$$

Dla oszacowania wariancji s^2 możemy skorzystać z rozkładu Gaussa jako modelu rozkładu błędów przypadkowych i obliczyć:

$$\mathcal{E}((x_i - \mu)^4) = 3\sigma^4,$$

(dla wszystkich wyników pomiarów x_i zakładamy ten sam rozkład błędów przypadkowych, a więc wartości $\mathcal{E}((x_i - \mu)^4)$ są dla wszystkich i takie same) co prowadzi do oceny (przybliżenia):

$$\text{Var}(s^2) = \mathcal{E}((s^2 - \sigma^2)^2) \approx \frac{2}{n-1} \sigma^4.$$

Zgodnie z przyjętą umową, jako ocenę σ^2 mamy (podana jest niepewność standardowa):

$$s^2 \pm \sigma^2 \sqrt{\frac{2}{n-1}}$$

Jako ocenę, jakiej to odpowiada zmienności samego s rozwiążmy względem δ poniższe równanie:

$$\frac{1}{2} ((\sigma + \delta)^2 - (\sigma - \delta)^2) = \sigma^2 \sqrt{\frac{2}{n-1}} \Rightarrow \delta = \frac{\sigma}{\sqrt{2(n-1)}}.$$

Tak otrzymana wartość δ jest miarą *niepewności* z jaką s przybliża wartość σ (do takiej samej oceny prowadzi dyskutowany w następnych wykładach wzór opisujący „propagację małych niepewności”). Tabela 2 w Uwadze 2 do §14 *Slajdów/transparencji do wykładu* ilustruje tę zależność i jednocześnie uzasadnia dlaczego wystarczy zaokrąglenie niepewności do dwóch cyfr znaczących – jak widać, dopiero dla $n > 52$ dokładność z jaką s przybliża σ jest lepsza niż 10%.

Przedstawiona powyżej i w §§12-14 *Slajdów/transparencji do wykładu* metoda wyznaczania niepewności pomiaru na podstawie serii niezależnych pomiarów – w języku statystyki: na podstawie pobranej próby zmiennej losowej – to tak zwana **metoda A** wyznaczania niepewności. Polega ona na wyznaczaniu parametrów rozkładów zmiennych losowych wyłącznie na podstawie pobranej próby (wyników serii pomiarów).

Uwzględnienie dokładności (rozdzielczości) przyrządów

Z dotychczasowej analizy można by wyciągnąć wniosek, że jeśli wyeliminowane są błędy grube i systematyczne, to wydłużenie serii pomiarów pozwoliłoby dowolnie zmniejszyć niepewność wyniku. Przecież, jak pokazaliśmy wcześniej, gdy $n \rightarrow \infty$ to:

$$n \rightarrow \infty \Rightarrow s^2 \rightarrow \sigma^2 \Rightarrow s_x^2 = \frac{s^2}{n} \rightarrow 0.$$

Należy jednak pamiętać, że pomiary wykonujemy za pomocą przyrządów o skończonej dokładności odczytu wyniku. Najłatwiej dostrzec to, gdy używamy przyrządu cyfrowego. Otrzymywane wyniki ograniczone są do kilku cyfr i znany jest najmniejszy krok (rozdzielczość przyrządu) o jaki może zmienić się wskazanie takiego przyrządu. Najczęściej ostatnia cyfra zmienia się co 1, 2 lub 5. Definiuje to **rozdzielczość przyrządu Δ** . Przykładem może być cyfrowy zegar na dworcu kolejowym – pokazuje on godziny i minuty, przy czym zawsze pokazuje czas, który „już minął”. Dokładny czas taki zegar wskazuje jedynie w chwili zmiany liczby minut. W każdym innym momencie późni się – średnio o pół minuty. Na dworcu ma to oczywiste zalety – pociągi odjeżdżają po upływie czasu zapisanego w rozkładzie. Zwykle jednak chcemy, żeby odczytana wartość była jak najbliższa wyniku dokładnego.

Oczekujemy, że przyrząd pomiarowy jest dokładny w tym sensie, że średni błąd wskazań jest równy zeru. Przełączania wskazań powinny więc następować w połowie odległości między kolejnymi wskazywanymi wynikami. Ideę działania tak wykalibrowanego przyrządu przedstawia Rysunek 5 *Slajdów/transparencji do wykładu*. Dla idealnie dokładnego miernika wykresem *odczytu wyniku* jako funkcji wartości wielkości mierzonej, czyli *sygnału na wejściu* byłaby cienka, niebieska prosta na Rysunku 5. Rzeczywisty odczyt wyniku, tj. *sygnał na wyjściu* przedstawiają poziome, zielone odcinki.

Odpowiedzią przyrządu cyfrowego jest zielona krzywa schodkowa.

Rysunek 6 *Slajdów/transparencji do wykładu* przedstawia wykres **błędu odczytu ξ** :

$$\text{błąd odczytu} := \xi = \text{odczyt wyniku} - \text{sygnał na wejściu}$$

Jak widać, błąd odczytu może być, z jednakowym prawdopodobieństwem, dowolną liczbą z zakresu $-a < \xi < a$, przy czym, w omawianym modelu $a = \Delta/2$. Rozkładem prawdopodobieństwa opisującym błąd przypadkowy jest więc **rozkład jednostajny** o gęstości:

$$g(\xi) = \frac{1}{2a},$$

gdy $-a \leq \xi \leq a$ oraz $g(\xi) = 0$ poza tym przedziałem. Wartość oczekiwana

$$\mathcal{E}(\xi) = \int_{-a}^a \xi g(\xi) = 0,$$

a wariancja:

$$\text{Var}(\xi) = \int_{-a}^a \xi^2 g(\xi) = \frac{1}{2a} \int_{-a}^a \xi^2 = \frac{a^2}{3}.$$

Instrukcja użycia przyrządu zawiera (w każdym razie powinna zawierać) dokładne wskazówki jaką liczbę należy przyjąć za wartość a . Zwykle jest to liczba większa niż połowa rozdzielczości odczytu i może zależeć od wartości wskazań. Informacje producenta dotyczące dokładności przyrządu i warunków jego prawidłowej pracy należy zawsze uwzględniać w planowaniu i przygotowaniu pomiarów i analizie ich wyników.

Tutaj, dla uproszczenia modelu i uniknięcia zbyt optymistycznej oceny dokładności pomiaru, przyjmijmy $a = \Delta$. Przyjmujemy ten sam model matematyczny opisu błędu odczytu w przypadku pomiaru przyrządem analogowym (np. linijką z podziałką milimetrową, wagą uchylną z podziałką co 1 dkg itp.) z podziałką o najmniejszej działce równej Δ .

Błąd odczytu jest błędem systematycznym, bo w powtarzanych pomiarach tej samej wielkości fizycznej (tego samego sygnału na wejściu przyrządu) ma tę samą wartość. Wartości tej jednak nie znamy. Wiemy tylko w jakim zakresie każda wartość błędu odczytu jest jednakowo prawdopodobna.

Wróćmy do analizy całości zagadnienia dokładności pomiaru (§16 *Slajdów/transparencji do wykładu*). Sygnał na wejściu przyrządu zawiera błąd przypadkowy ε , a z odczytem jego wartości jest związany błąd odczytu ξ . Każdy z tych błędów jest zmienną losową i są to dwie różne, niezależne zmienne losowe. Wartość oczekiwana sumy niezależnych zmiennych losowych równa się sumie ich wartości oczekiwanych:

$$\mathcal{E}(\varepsilon + \xi) = \mathcal{E}(\varepsilon) + \mathcal{E}(\xi) = 0,$$

a wariancja sumy niezależnych zmiennych losowych równa jest sumie ich wariancji:

$$Var(\varepsilon + \xi) = Var(\varepsilon) + Var(\xi) = s_x^2 + \frac{\Delta^2}{3}$$

– łatwo to wykazać odwołując się do definicji niezależnych zmiennych losowych. Końcową **niepewność standardową** $u(\bar{x})$ wyniku bezpośredniego pomiaru pojedynczej wielkości fizycznej jest więc równa:

$$u(\bar{x}) = \sqrt{s_x^2 + \frac{\Delta^2}{3}}.$$

Tak otrzymany wynik: *wartość i jej niepewność* zapisujemy zgodnie z regułami sformułowanymi w §14 *Slajdów/transparencji do wykładu*.

Także i w tym przypadku należy określić sposób wyznaczenia liczby podawanej po zanku „ $\pm$ ”.

Przedstawiony w tym paragrafie sposób wyznaczania niepewności – połączenie obliczeń na podstawie wyników losowej próby (serii pomiarów) i modelu matematycznego (w tym przypadku opisującego dokładność odczytu wskazań przyrządu), to tzw. **metoda B** wyznaczania niepewności wyniku pomiaru.

Przyjęte międzynarodowe standardy obliczania i wyrażania niepewności pomiaru zostały szczegółowo przedstawione na stronie Głównego Urzędu Miar:

https://gum.gov.pl/ftp/pdf/Publikacje/Przewodnik_JCGM_100_ver__fin_27_08_2019_popr_.pdf

Jest to tłumaczenie dokumentu opublikowanego na stronie BIPM (Międzynarodowego Biura Miar i Wag w Paryżu).

Zainteresowanych odsyłam do tej strony. Równoważne zalecenia można znaleźć także na stronie amerykańskiego NIST (National Institute of Standards and Technology):

<https://www.nist.gov/pml/nist-technical-note-1297>

4 Wprowadzenie do testowania hipotez – test „ 3σ ”

Do tej pory rachunek prawdopodobieństwa wykorzystywaliśmy jedynie do obliczania wartości oczekiwanych zmiennych losowych. O pojawiających się rozkładach prawdopodobieństwa zakładaliśmy niewiele poza tym, że potrzebne nam wartości oczekiwane istnieją i są skończone. Pozwoliło nam to znajdować oceny (estymatory) mierzonych wielkości fizycznych. Dokładna znajomość prawdopodobieństw zdarzeń jest za to niezbędna, gdy na podstawie doświadczeń (pomiarów) **poddajemy testom hipotezy przyjmowane podczas badań**.

Na przykład:

- Rozważania teoretyczne prowadzą do wniosku, że pewna wielkość ma wartość μ – to nasza hipoteza.
- Wykonujemy pomiary i otrzymujemy jako ich wynik wartość $\hat{x}$ i jej niepewność $u(\hat{x})$, przy czym, na ogół $\hat{x} \neq \mu$.
- Pytanie: Jak ocenić, czy nasz wynik jest zgodny z przewidywaniami teorii? Jak duża różnica między $\hat{x}$ i μ upoważnia do odrzucenia hipotezy, tzn. odrzucenia modelu teoretycznego?

Sposób formułowania i rozstrzygania tego pytania w ramach statystyki ilustruje Przykład 8 w §18 *Slajdów/transparencji do wykładu*. Wyobraźmy sobie, że znajomy, dla zabicia jego i naszej nudy, zaproponował nam grę w „orła i reszkę”. Dodatkowo, żeby gra była emocjonująca zaproponował, że za każdym razem, kiedy wypadnie *reszka* on płaci nam k złotych, a gdy wypadnie *orzeł*, my płacimy jemu taką samą stawkę. Przebieg gry w pierwszych 8 rzutach opisuje Przykład 8: 7 razy wypadł *orzeł*. Czy powinniśmy uznać, że moneta znajomego nie jest „uczciwa” i przerwać grę? Gdy moneta jest uczciwa taki wynik nie jest niemożliwy, ale wydaje się nam mało prawdopodobny – rachunek prawdopodobieństwa pozwala obliczyć, że rzucając uczciwą monetą, taki wynik (7 lub więcej orłów w 8 rzutach) zdarzałby się rzadziej niż raz na 28 rozgrywek. Nasza decyzja będzie zależała od tego, jak bardzo bolesna jest dla nas strata k złotych – to możliwy koszt decyzji „gram dalej”. Z drugiej strony, przerwanie gry i zarzucenie znajomemu nieuczciwości też może być kosztowne, zwłaszcza jeśli jest to np. nasz szef. Musimy zdecydować jakie ryzyko przerwania gry, gdy moneta jest uczciwa jesteśmy w stanie zaakceptować. W języku matematyki: jakie prawdopodobieństwo odrzucenia hipotezy prawdziwej (tj. **popęlnienia błędu I rodzaju**) jesteśmy skłonni zaakceptować.

W Przykładzie 8, *testowana hipoteza*, to stwierdzenie *moneta jest uczciwa*. Prawdopodobieństwo wyniku rzeczywiście zaobserwowanego lub jeszcze mniej korzystnego dla testowanej hipotezy, obliczone zostało przy założeniu, że testowana hipoteza jest *prawdziwa*. Jeżeli tak obliczone prawdopodobieństwo jest mniejsze od przyjętej przez nas *wartości krytycznej* – akceptowanego prawdopodobieństwa odrzucenia hipotezy prawdziwej, to hipotezę odrzucamy, tzn. uznajemy, że jest *fałszywa*. W przeciwnym razie *nie odrzucamy testowanej hipotezy* – nie znaczy to, że uznajemy ją za prawdziwą, a jedynie, że w świetle zebranych danych (wyników serii rzutów)

nie mamy podstaw do jej odrzucenia.

W przypadku porównania wyniku pomiaru z wartością przewidywaną teoretycznie (Pytanie 1) postępowanie jest dokładnie analogiczne. Jeżeli założymy, że wynik naszego pomiaru podlega rozkładowi Gaussa, to znając parametry tego rozkładu potrafimy obliczyć prawdopodobieństwo tego, że różni się on od zadanej (przewidzianej teoretycznie) wartości więcej niż obserwowana rzeczywiście różnica. Obliczone prawdopodobieństwo porównujemy z przyjętą przez nas wartością krytyczną. W „codziennej” praktyce za tę wartość krytyczną przyjmuje się zwykle prawdopodobieństwo odpowiadające różnicy wyniku pomiaru i przewidywań teorii większej niż trzy razy niepewność pomiaru – odpowiada to wartości $3 \cdot \sigma$ rozkładu Gaussa i stąd nazwa testu.

Pytanie 2 z §17 *Slajdów/transparencji do wykładu* dotyczy zgodności wyników dwóch pomiarów tej samej wielkości (testowana hipoteza), ale z różnymi niepewnościami. Udzielając odpowiedzi korzystamy tu z faktu, że różnica dwóch zmiennych losowych podlegających rozkładowi Gaussa o takich samych wartościach oczekiwanych, ale różnych wariancjach, podlega rozkładowi Gaussa o wartości oczekiwanej zero i wariancji równej sumie wariancji odejmowanych zmiennych.

Prawdopodobieństwo krytyczne odpowiadające różnicy „ 3σ ” jest tu wyborem arbitralnym, ale dosyć częstym.

Przykładowe zadania do omówionej powyżej części wykładu.

Zadanie 1. Wykonano $n = 245$ pomiarów pojedynczego okresu drgań wahadła matematycznego. Jeśli przez $T_i, i = 1, \dots, n$ oznaczmy wynik pomiaru i -tego okresu, to:

$$\sum_{i=1}^n T_i = 694,40 \text{ s}, \quad \sum_{i=1}^n (T_i - \bar{T})^2 = 1,1824 \text{ s}^2.$$

Zakładając, że pomiary te wykonano stoperem pozwalającym na odczyt czasu z dokładnością $\Delta = 0,01$ s wyznacz okres wahadła i jego niepewność.

Zadanie 2. Podczas zajęć pracowni fizycznej pięciu kolejnych studentów mierzyło, każdy jednokrotnie, wartość współczynnika załamania n tej samej szklanej płytki płasko-równoległościennej. Pomiary wykonywali metodą pomiaru przesunięcia promienia świetlnego, jakiemu ulega on po przejściu przez płytkę. Każdy ze studentów wykonał pomiar dokładnie w ten sam sposób, rzucając promień świetlny o tej samej długości fali na płytkę pod tym samym kątem. Otrzymali oni następujące wartości współczynnika załamania: 1,4629; 1,4392; 1,4761; 1,5011 oraz 1,4992. Niestety, ostatni student po wykonaniu pomiaru stłukł płytkę. Asystent wiedział, że płytka wykonana była z jednego z typów szkła:

- kwarcowego o współczynniku załamania 1,4584;
- kron o współczynniku załamania 1,5162;
- flint kron o współczynniku załamania 1,6032 lub
- szkła ołowiowego kron o współczynniku załamania 1,7550.

Płytkę z jakiego typu szkła musi zamówić asystent, aby uzupełnić brak?

Zadanie 3. Podczas zajęć pracowni fizycznej trzech studenci: A, B oraz C mierzyli gęstość trzech próbek metalu (każdy ze studentów badał jedną próbkę) dostarczonych przez asystenta. O asystencie wiadomo, że dysponuje tylko dwoma typami próbek, jakie daje studentom do analizy i są to próbki żelaza i chromu. W tablicach można znaleźć, że gęstość żelaza wynosi $7,874 \text{ g/cm}^3$, a chromu $7,140 \text{ g/cm}^3$. Dwaj studenci A i B wykonali prostsze pomiary i z momentem zakończenia zajęć otrzymali następujące rezultaty: $\rho_A = (7,09 \pm 0,15) \text{ g/cm}^3$, $\rho_B = (8,01 \pm 0,15) \text{ g/cm}^3$. Trzeci student wybrał bardziej pracochłonną metodę i w momencie zakończenia zajęć dysponował następującym czterokrotnym pomiarem gęstości swojej próbki: 7,03, 6,67, 7,23 oraz 7,61 (wszystkie wyniki w g/cm^3).

- (a) Przeprowadź analizę, jaką będzie musiał wykonać student C i podaj ocenę gęstości ρ_C oraz jej niepewność, jakie wynikają z jego danych.
- (b) Zidentyfikuj rodzaj metalu, który badał każdy ze studentów.