Wojciech TOCZEK
Politechnika Gdańska
Wydział Elektroniki, Telekomunikacji i Informatyki

WYRAŻANIE NIEPEWNOŚCI ZA POMOCĄ PRZEDZIAŁÓW

Z perspektywy dwóch różnych interpretacji prawdopodobieństwa - klasycznej (częstościowej) i subiektywnej (bayesowskiej) oraz propozycji nowego przewodnika ustalającego zasady obliczania i wyrażania niepewności pomiaru (GUM), porównano sposoby komunikowania niepewności za pomocą przedziałów: ufności, bayesowskiego, objęcia, rozszerzenia.

Słowa kluczowe: niepewność pomiaru, przedział objęcia, przedział bayesowski,

EXPRESSION OF UNCERTAINTY IN MEASUREMENT USING INTERVALS

From the perspective of two different interpretations of probability - objective (frequentist) and subjective (Bayesian) as well as the proposition of the revised Guide to the Expression of Uncertainty in Measurement (GUM), ways of announce of uncertainty using: confidence interval, Bayesian interval, coverage interval and expanded interval are compared.

Keywords: uncertainty of measurement, coverage interval, Bayesian interval,

1. WPROWADZENIE

Wynik pomiaru można podawać w formie estymaty menzurandu i jej niepewności standardowej lub w formie przedziału, który zawiera wartość menzurandu z określonym prawdopodobieństwem. Druga forma jest przydatna przy podejmowaniu decyzji, na przykład w trakcie badania zgodności.

W referacie porównano sposoby komunikowania niepewności pomiaru za pomocą przedziałów, z perspektywy dwóch różnych interpretacji prawdopodobieństwa - klasycznej (tzw. częstościowej) i subiektywnej (prawdopodobieństwo jako miara przekonania). Celem referatu jest wyjaśnienie zróżnicowania i relacji między czterema wybranymi przedziałami: ufności, bayesowskim, objęcia i rozszerzenia. Przedstawiona analiza pomaga głębiej zrozumieć koncepcję propozycji nowego przewodnika GUM [1], w którym preferowana jest subiektywna (bayesowska) interpretacja prawdopodobieństwa.

2. DWA SPOSOBY WYRAŻANIA NIEPEWNOŚCI

Kluczem do zrozumienia różnic w sposobach komunikowania niepewności jest rozróżnienie dwóch sposobów interpretowania prawdopodobieństwa – klasycznego i bayesowskiego.

W klasycznej interpretacji, prawdopodobieństwo jest częstością występowania zdarzenia w serii identycznych prób. Definicja ta nie ma zastosowania do zdarzeń niepowtarzalnych.

Bayesowskie rozumienie prawdopodobieństwa, to miara racjonalnego zaufania w prawdziwość danej tezy, uwarunkowana posiadaną informacją. Takie ujęcie ma zastosowanie zarówno do zdarzeń powtarzalnych jak i niepowtarzalnych. Przypadkowość rozumiana jest jako wyraz niepełnej informacji jaką posiadamy o tezie.

Różne pojmowanie prawdopodobieństwa prowadzi do dwóch sposobów wyrażania niepewności.

2.1. Wyrażanie niepewności w częstościowej koncepcji prawdopodobieństwa – przedział ufności

W ujęciu częstościowym, wielkości estymowane za pomocą danych pomiarowych są stałymi parametrami statystycznymi o nieznanej wartości. Pojęcie **przedziału ufności** (*confidence interval*) dotyczy prawdopodobieństwa wzajemnego położenia uzyskanego oszacowania nieznanego parametru i jego prawdziwej wartości. Przedział ufności jest realizacją losowego przedziału, który przy powtarzaniu prób z tej samej populacji obejmie wartość wielkości estymowanej w $100\,(1-\alpha)$ % wszystkich prób. Przedział buduje się wokół wartości średniej, otrzymanej z próby. **Poziom ufności** jest zwyczajowo zapisywany jako $(1-\alpha)$, gdzie α jest małą nieujemną liczbą i jest podawany w procentach $100\,(1-\alpha)$ %. Przedział ufności obliczony z kolejnych prób będzie na ogół różny. Ponadto niektóre z wyznaczonych przedziałów będą zawierały prawdziwą wartość parametru, a inne nie będą jej zawierały. Jeżeli wynikiem analizy jest 100 przedziałów ufności wyznaczonych z kolejnych prób dla poziomu ufności 95 %, to średnio 95 na 100 przedziałów będzie zawierało wartość parametru populacji. Nie ma tu mowy o prawdopodobieństwie, że nieznana wartość parametru będzie zawarta w jakimś stałym przedziale, ponieważ nieznany parametr nie jest zmienną losową. Gdy liczność próby n wzrasta, to przedział ufności zawęża się. Dla $n \to \infty$ szerokość przedziału ufności zdąża do 0.

2.2. Wyrażanie niepewności w ujęciu subiektywnym – przedział bayesowski

W przeciwieństwie do podejścia częstościowego, w ujęciu subiektywnym, estymowany parametr jest zmienną losową, a wyniki pomiarów są wartościami ustalonymi. **Przedział bayesowski** (*Bayesian interval* lub *credible interval*), konstruowany z rozkładu a posteriori, zawiera wartość parametru θ, z prawdopodobieństwem uwarunkowanym dostępnymi danymi pomiarowymi $\boldsymbol{y}$. Jest to prawdopodobieństwo subiektywne, ponieważ zależy od wykorzystanej wiedzy o menzurandzie zakodowanej w postaci funkcji gęstości prawdopodobieństwa a priori.

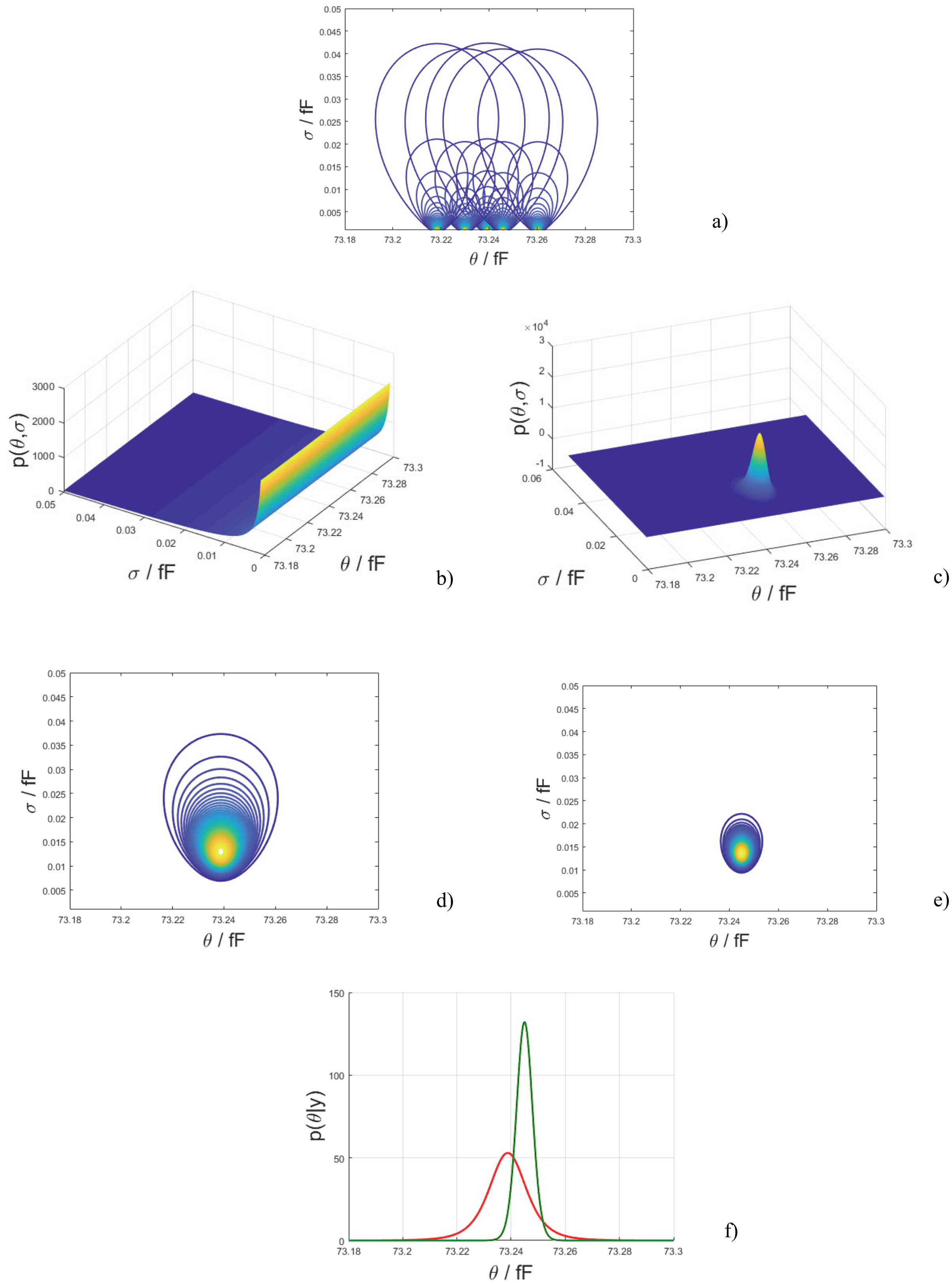
Rozkład a posteriori, reprezentujący wiedzę o menzurandzie po wykonaniu serii pomiarów, wyznaczamy za pomocą twierdzenia Bayesa (Rys. 1). Wyniki pomiarów podstawiamy sukcesywnie do funkcji likelihood, która ma postać jednakową dla każdego wyniku (Rys. 1a). Funkcja likelihood nie zmienia kształtu, jest jedynie przesuwana w dziedzinie wielkości mierzonej [2]. Kształt rozkładu a posteriori zależy od użytego rozkładu a priori. Gdy nie posiadamy wiedzy o menzurandzie, jako rozkład a priori stosujemy iloczyn rozkładu równomiernego (dla nieznanej wartości oczekiwanej) i rozkładu o postaci σ^{-1} (dla nieznanej wartości odchylenia standardowego), tzw. rozkład nieinformacyjny (Rys. 1b), otrzymując rozkład a posteriori przedstawiony na Rys. 1d. Wykorzystując wiedzę teoretyczną i dodatkową informację spoza aktualnego pomiaru, możemy skonstruować informacyjny rozkład a priori (Rys. 1c), który doprowadza rozkład a posteriori do bardziej wysmukłej postaci, przedstawionej na Rys. 1e. Uniezależnienie rozkładów a posteriori od σ, za pomocą marginalizacji, prowadzi do krzywych przedstawionych na Rys. 1f.

Przedział bayesowski $C=(a,b)$ wyznaczamy numerycznie ze zmarginalizowanego rozkładu a posteriori. Znajdujemy a i b takie, że

$$\int_{-\infty}^{a} p(\theta|\boldsymbol{y})\,d\theta = \int_{b}^{+\infty} p(\theta|\boldsymbol{y})d\theta = \frac{\alpha}{2}. \tag{1}$$

Wtedy

$$P(\theta \in C|\boldsymbol{y}) = \int_{a}^{b} p(\theta|\boldsymbol{y})d\theta = 1-\alpha. \tag{2}$$



Rys. 1. Przykład rozkładów a posteriori uzyskanych na bazie danych pomiarowych, a) kontury funkcji likelihood dla 5 wyników pomiarów, b) nieinformacyjny rozkład a priori, c) informacyjny rozkład a priori, d, e) odpowiednie rozkłady a posteriori, f) zmarginalizowane rozkłady a posteriori: niższy dla nieinformacyjnego rozkładu a priori, wyższy dla informacyjnego

3. KOMUNIKOWANIE NIEPEWNOŚCI WEDŁUG PROPOZYCJI NOWEGO GUM

W metrologii, zamiast *przedziału ufności* od dawna stosowane jest pojęcie *coverage interval*, co można tłumaczyć jako **przedział objęcia**. Nieadekwatność terminu *przedział ufności* wynika stąd, że procedura obliczania standardowej niepewności pomiaru jest bardziej rozbudowana od statystycznej procedury wyznaczania przedziału ufności. Skrócony przepis obliczania standardowej niepewności pomiarowej w pomiarach pośrednich, obejmuje następujące kroki:
1. Wyrażenie związku między wielkością mierzoną Y i wielkościami X_i, od których Y zależy, w postaci funkcji $Y = f(X_1, X_2, \dots, X_n)$.
2. Wyznaczenie estymat x_i i ich niepewności standardowych $u(x_i)$ tych wartości wielkości wejściowych, dla których można zastosować metodę typu A.
3. Wyznaczenie estymat x_i i ich niepewności standardowych $u(x_i)$ pozostałych wartości wielkości wejściowych metodą typu B.
4. Obliczenie wyniku pomiaru, to jest estymaty y wielkości mierzonej Y z zależności funkcyjnej f, dla wartości wielkości wejściowych X_i równych estymatom x_i .
5. Określenie niepewności standardowej $u(y)$ wyniku pomiaru y na podstawie niepewności standardowych związanych z estymatami wartości wielkości wejściowych.

Dla danej serii pomiarowej i prawdopodobieństwa p, końce przedziału ufności i przedziału objęcia będą różne. Uzasadnia to stosowanie dwóch różnych nazw dla omawianych przedziałów.

W propozycji nowego przewodnika GUM [1] zadeklarowano przyjęcie bayesowskiej interpretacji prawdopodobieństwa, pozostając jednak przy dotychczasowym nazewnictwie przedziału komunikującego niepewność. W [1] znajdujemy definicję przedziału objęcia „Przedział, który z określonym prawdopodobieństwem, bazującym na dostępnej informacji, zawiera wartość menzurandu”. Prawdopodobieństwo p, że wartość menzurandu znajduje się w przedziale objęcia jest nazywane **prawdopodobieństwem objęcia** (*coverage probability*).

Według [1] sposób konstruowania przedziału objęcia zależy od posiadanej wiedzy o menzurandzie. Proponowane postępowanie obejmuje także stan braku wiedzy. W przypadku, gdy rozkład prawdopodobieństwa menzurandu nie jest znany, zalecane jest wyznaczanie „distribution-free coverage intervals” z użyciem nierówności Czebyszewa

$$P\{\lceil y - Y \rceil \leq k_p u(y)\} \geq 1 - \frac{1}{k_p^2}, \tag{3}$$

gdzie: y jest najlepszą estymatą menzurandu Y.
Dla założonego prawdopodobieństwa objęcia p, prawa strona nierówności (3) pozwala obliczyć współczynnik objęcia $k_p = 1/(1-p)^{\frac{1}{2}}$. Na tej podstawie konstruuje się przedział $y \pm U_p$, gdzie $U_p = k_p u(y)$. Dla $p = 0{,}95$ przedział objęcia wynosi

$$y \pm 4{,}47u(y). \tag{4}$$

Zdefiniowany powyżej przedział objęcia ma tę zaletę, że nie wymaga założeń odnośnie rozkładu prawdopodobieństwa menzurandu. Gdy wiadomo, że rozkład jest jednomodalny i symetryczny, to przedział objęcia (4) można zawęzić korzystając z nierówności Gaussa

$$P\{\lceil y - Y \rceil \leq k_p u(y)\} \geq 1 - \frac{4}{9}\frac{1}{k_p^2}. \tag{5}$$

Z nierówności (5) dostajemy współczynnik objęcia $k_p = 2/\left[3(1-p)^{\frac{1}{2}}\right]$. Stąd, dla $p = 0{,}95$ przedział objęcia wynosi

$$y \pm 2{,}98u(y). \tag{6}$$

W przypadku, gdy dysponujemy funkcją gęstości prawdopodobieństwa menzurandu, przedział objęcia można wyznaczyć numerycznie, tak aby spełniał wzór (1). Dla kilku typów rozkładów dostępne są w literaturze relacje pomiędzy p i k_p. Na przykład dla rozkładu Gaussa

$$p = \frac{1}{\sqrt{2\pi}} \int_{-k_p}^{k_p} e^{-\frac{x^2}{2}} dx\,. \tag{7}$$

Dla współczynnika $k_p = 1{,}96$ otrzymujemy z (7) prawdopodobieństwo objęcia $p = 95\ \%$.

Dla asymetrycznych rozkładów prawdopodobieństwa przedział objęcia może być podawany w formie (y_a, y_b), gdzie y_a i y_b wyznaczają przedział na osi odciętych, taki, że pole powierzchni pod krzywą rozkładu, rozpiętą nad przedziałem, wynosi p. Dla danego prawdopodobieństwa p, jest nieskończenie wiele takich przedziałów. Wybór zależy od rodzaju zastosowania wyniku pomiaru [1].

Niepewność objęcia (*coverage uncertainty*) jest bardziej ogólną koncepcją od **niepewności rozszerzonej** (*expanded uncertainty*), stosowanej w dotychczasowej wersji przewodnika GUM. Z definicji *lewa i prawa niepewność objęcia*, to odległości od dolnej i górnej granicy przedziału objęcia do wartości odpowiadającej estymacie menzurandu. Niepewność objęcia odpowiada niepewności rozszerzonej, gdy rozkład wartości menzurandu jest symetryczny [3].

4. PODSUMOWANIE

W ujęciu bayesowskim, estymowany parametr jest zmienną losową i możemy na przykład stwierdzić „Wartość menzurandu leży w przedziale (a, b) z prawdopodobieństwem 0,95 wyznaczonym na podstawie dostępnej informacji". W ujęciu klasycznym takie stwierdzenie nie ma sensu, ponieważ estymowany parametr jest stały, chociaż nieznany i leży w wyznaczonym przedziale lub nie leży. Pomimo tego przedział ufności, związany z klasyczną interpretacją prawdopodobieństwa, jest często mylnie interpretowany w sposób właściwy przedziałowi bayesowskiemu. Jednak podejście częstościowe jest użyteczne, gdy wielokrotnie wykonujemy eksperyment i wielokrotnie konstruujemy przedział ufności. Otrzymujemy wtedy zbiór przedziałów ufności, z których $100\ (1-\alpha)\ \%$ zawiera wartość menzurandu. Poziom ufności dotyczy przedziałów a nie wartości menzurandu.

Przewodnik [1] daje wskazówki jak określić przedział objęcia, który zawiera wartość menzurandu z określonym prawdopodobieństwem, rozumianym jako miara przekonania. A więc do oceny niepewności zaleca podejście bayesowskie. W [1] brak jest rozwinięcia tematu dotyczącego wykorzystania dodatkowej wiedzy o menzurandzie i przedstawienia aparatu statystyki bayesowskiej, którego można by użyć w celu zawężenia przedziału objęcia.

LITERATURA

1. Joint Committee for Guides in Metrology: JCGM 201X Evaluation of measurement data – Guide to uncertainty in measurement. Committee draft.
2. Toczek W.: Zastosowanie statystyki bayesowskiej do uzasadnienia zmiany sposobu obliczania standardowej niepewności pomiaru. Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej, 57, 2017.
3. Lira I.: Evaluating the measurement uncertainty. Fundamentals and practical guidance. Institute of Physics Publishing, 2002.