

# **Nowa, kompletnościowa interpretacja prawdopodobieństwa**

Wykład podstawowy

**Andrzej Piegat**

Wydział Informatyki,

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie,

Żołnierska 49, 71-062 Szczecin

apiegat@wi.zut.edu.pl, <http://apiegat.strony.wi.ps.pl>

## **Streszczenie**

Wykład przedstawia nową (według wiedzy autora) interpretację prawdopodobieństwa, która nie posiada szeregu wad powszechnie stosowanej interpretacji częstościowej. Wykład pokazuje, że interpretacja częstościowa nie jest wiarygodna w przypadku braku danych oraz małej i średniej liczby danych i tłumaczy przyczynę tego zjawiska. Przyczyną tą jest nieliniowy charakter tej interpretacji który powoduje, że dane o jednakowej wadze wpływają na prawdopodobieństwo w niejednakowy sposób. Fakt ograniczonej wiarygodności częstościowej interpretacji prawdopodobieństwa w omawianym przypadku nie świadczy jednak, że interpretacja ta jest całkowicie błędna, bowiem jej wiarygodność wzrasta z liczbą danych i przy odpowiedniej liczności danych jej wyniki stają się wiarygodne. Nowa, kompletnościowa interpretacja prawdopodobieństwa przedstawiona w wykładzie dostarcza wiarygodne wyniki zarówno w przypadku braku danych jak i w przypadku małej, średniej, i dużej liczby danych. Interpretacja ta ma charakter liniowy w tym sensie, że dane o jednakowej wadze w jednakowy sposób wpływają tu na prawdopodobieństwo hipotezy. Zakres jej dokładności jest więc większy niż zakres interpretacji częstościowej.

## **Słowa kluczowe**

Prawdopodobieństwo, teoria prawdopodobieństwa, statystyka, interpretacje prawdopodobieństwa, kompletnościowa interpretacja prawdopodobieństwa, częstościowa interpretacja prawdopodobieństwa, niepewność, teoria niepewności, modelowanie niepewności.

# Spis treści

|   |   |    |
|---|---|----|
| 1 | Wprowadzenie  | 3  |
| 2 | Klasyczna i częstościowa interpretacja prawdopodobieństwa   | 4  |
| 3 | Główne zarzuty względem klasycznej i częstościowej interpretacji prawdopodobieństwa                           | 5  |
| 4 | Nowa, kompletnościowa interpretacja prawdopodobieństwa  | 11 |
| 5 | Niepewność prawdopodobieństwa   | 17 |
| 6 | Optymalna reprezentacja $p_{hR}$ zakresu niepewności $[p_{hmin}, p_{hmax}]$ prawdopodobieństwa $p_h$ hipotezy | 24 |
| 7 | Podsumowanie  | 32 |
|   | Literatura  | 33 |

# 1 Wprowadzenie

Teoria prawdopodobieństwa jest chyba najstarszą, choć nie jedyną (systemy rozmyte Zadeha [12, 21], teoria wiary Dempstera-Schafera [19], teoria możliwości Duboisa i Prade [4], teoria luk informacyjnych Yakova [20]).

Pierwszą interpretację prawdopodobieństwa podał Laplace [13] w 1814 roku (interpretacja klasyczna). Od tego czasu wydano tysiące książek i artykułów na temat prawdopodobieństwa. Teoria ta jest nauką wykładaną na tysiącach uczelni i powszechnie stosowaną w praktyce, np. w badaniach statystycznych. Stąd wydawać by się mogło, że jest nauką silnie ugruntowaną, nauką o silnych podstawach, nie budzącą żadnych wątpliwości. Okazuje się jednak, że prawda jest zupełnie inna. Niepewność co do tego czym jest prawdopodobieństwo jest obecnie bardzo duża. Sformułowano bowiem wiele pytań, wątpliwości i przedstawiono wiele paradoksów wynikających z obecnego rozumienia (interpretacji) prawdopodobieństwa. Z tego względu niektórzy naukowcy uważają, że obecna teoria prawdopodobieństwa poniosła przynajmniej częściową klęskę. Przykładem takiej opinii jest wydana w 2009 roku książka o wymownym tytule „The search for certainty. On the clash of philosophy of probability” napisana przez profesora University of Washington Krzysztofa Burdzy [1]. Książka ta wzbudziła ożywioną dyskusję prowadzoną głównie na stronach internetowych [6, 10]. Niektórzy naukowcy gorąco popierają opinię profesora Burdzy, inni się jej sprzeciwiają, chociaż raczej w sposób umiarkowany. Powołują się przy tym na fakt praktycznych korzyści ze stosowania obecnej teorii prawdopodobieństwa w statystyce. Jednak poglądy prof. Burdzy nie są odosobnione. W literaturze przedmiotu można znaleźć wiele przykładów i paradoksów pokazujących słabe strony obecnej teorii prawdopodobieństwa. Spotyka się nawet opinie krańcowe jak: „probability does not exist” (prawdopodobieństwo nie istnieje), „no matter how much information you have, there is no scientific method to assign a probability to an event” (niezależnie od tego ile posiadasz informacji, nie istnieje żadna naukowa metoda przyporządkowania prawdopodobieństwa zdarzeniu), wypowiedziana przez sławnego probabilistę de Finettiego w [5]. Ze względu na chęć ograniczenia objętości niniejszego wykładu nie można przedstawić w nim wszystkich krytycznych uwag dotyczących współczesnej teorii prawdopodobieństwa. Jednak zainteresowany czytelnik łatwo może je znaleźć w książce prof. Burdzy [1], w książkach innych autorów [11, 18], jak i na licznych stronach internetowych, np. [7, 8].

Przed wszystkim, zaskakującym jest, że mimo, iż żyjemy już w XXI wieku, to nie wiadomo jeszcze czym naprawdę jest prawdopodobieństwo, bowiem wśród naukowców nie ma na ten temat zgody choć intuicyjnie wydaje nam się, że rozumiemy jego istotę. Jednak w miarę zagłębiania się w problematykę prawdopodobieństwa każdy zaczyna stopniowo sobie uświadamiać jak trudny jest to temat. Istnieje wiele szkół naukowych interpretujących (wyjaśniających) prawdopodobieństwo w różny sposób. Poszczególne interpretacje prawdopodobieństwa starają się wypełnić braki innych interpretacji. Poniżej podane są główne interpretacje z krótkimi komentarzami prof. Burdzy [1]. Z interpretacjami prawdopodobieństwa można też zapoznać się w książkach Khrennikova [11] i Rocchiego [18] bądź na stronach internetowych [8] i w encyklopediach naukowych [7].

1. Klasyczna teoria prawdopodobieństwa (Laplace, 1814 [13]),  
„Według teorii klasycznej prawdopodobieństwo jest symetrią”.
2. Logiczna teoria prawdopodobieństwa (Carnap, 1950 [2]),  
„Według teorii logicznej prawdopodobieństwo jest ‚słabą’ implikacją”.

3. Częstościowa teoria prawdopodobieństwa (von Mises, 1957 [15]),  
„Według teorii częstościowej prawdopodobieństwo jest częstością w długim ciągu zdarzeń”.
4. Subiektywna teoria prawdopodobieństwa (de Finetti, 1975, [5]),  
„Według subiektywnej teorii prawdopodobieństwo jest opinią osoby”.
5. Skłonnościowa (propensity) teoria prawdopodobieństwa (Popper, 1957, [17]),  
„Według teorii skłonnościowej prawdopodobieństwo jest cechą fizykalną”.

Nie są to wszystkie interpretacje prawdopodobieństwa lecz najbardziej znane. Po szczególne interpretacje silnie się różnią i mają swoich zaciętych zwolenników. Istnieje też bardzo ciekawa, kompromisowa interpretacja prof. Hajek’a [7] w opinii którego każda z powyższych interpretacji jest częścią pewnej, złożonej prawdy o prawdopodobieństwie i pokazuje jedną z jego wielu twarzy. W praktyce najczęściej stosowanymi, wykładanymi na uczelniach, i dlatego najbardziej znanymi interpretacjami są: interpretacja klasyczna oraz częstościowa. Dlatego interpretacje te zostaną w skrócie przedstawione i skomentowane w następnym rozdziale.

## 2 Klasyczna i częstościowa interpretacja prawdopodobieństwa

Klasyczna interpretacja, której głównym twórcą jest Laplace [13] „przypisuje prawdopodobieństwo w sytuacji braku jakiegokolwiek materiału dowodowego (evidencji) bądź też na podstawie symetrycznie zrównoważonej ewidencji. Zakłada się tutaj równomierny rozdział prawdopodobieństwa pomiędzy wszystkie możliwe wyniki (co nie jest prawda w przypadku wielu problemów rzeczywistych – uwaga autora), tak więc klasyczne prawdopodobieństwo zdarzenia jest po prostu ułamkiem wynikającym z całkowitej liczby możliwości pod jakimi dane zdarzenie może zaistnieć” [7]. „Matematycznie można to sformułować w sposób następujący: jeśli przypadkowy eksperyment może zakończyć się  $N$  wzajemnie wykluczającymi się i równie prawdopodobnymi wynikami i jeśli  $N_A$  z tych wyników stanowi realizację zdarzenia  $A$ , to prawdopodobieństwo  $A$  jest zdefiniowane przez:

$$P(A) = \frac{N_A}{N}. \quad (1)$$

Definicja klasyczna posiada dwa główne ograniczenia. Po pierwsze, można ją stosować tylko w sytuacjach, w których istnieje ‚skończona’ liczba możliwych wyników. Ale pewne ważne probabilistyczne eksperymenty, takie jak rzucanie monetą do momentu pojawienia się reszki mogą mieć nieskończoną liczbę wyników. A po drugie, należy wcześniej sprawdzić, czy wszystkie możliwe wyniki są jednakowo prawdopodobne nie wykorzystując pojęcia prawdopodobieństwa w celu uniknięcia cyrkularności definicji – na przykład posługując się symetrią” [7]. Próby poprawienia klasycznej interpretacji i usunięcia chociażby niektórych jej wad podjęli się ‚frekwentyści’, których głównym przedstawicielem był von Mises [15]. „Frekwentyści twierdzą, że prawdopodobieństwo zdarzenia to jego względna częstość czasowa [7], to znaczy, względna częstość jego zachodzenia stwierdzona po powtórzeniu procesu wielką liczbą razy w podobnych warunkach . . . . Jeśli przez  $n_A$  oznaczymy liczbę zajścia zdarzenia  $A$  w  $n$  próbach, to jeśli zachodzi (2):

$$\lim_{n \rightarrow \infty} \frac{n_A}{n} = p, \quad (2)$$

mówimy, że  $P(A) = p$  [7]”.

Częstościowa interpretacja prawdopodobieństwa nazywana jest także interpretacją długiego ciągu zdarzeń (long-run frequency interpretation). W praktyce nie można jednak nigdy przeprowadzić niekończonej liczby eksperymentów. Z tego względu w praktyce stosowana jest tzw. częstościowa interpretacja skończonego ciągu zdarzeń (finite-frequency interpretation), w której opieramy się na takiej liczbie  $n$  eksperymentów jaką realnie mamy do dyspozycji [7]. Według tej interpretacji estymatę prawdopodobieństwa obliczamy ze wzoru (3):

$$p(A) = \frac{n_A}{n}, \quad (3)$$

gdzie:  $n$  – liczba skończona.

### 3 Główne zarzuty względem klasycznej i częstościowej interpretacji prawdopodobieństwa

Jak wspomniano w rozdziale 1 istnieje 5 głównych interpretacji prawdopodobieństwa i względem każdej z nich istnieją bardzo poważne zastrzeżenia. Ze względu na ograniczoną wykładu przedstawione zostaną niektóre zastrzeżenia tylko względem interpretacji klasycznej i częstościowej.

1.

”Ponieważ klasyczna definicja dotyczy tylko tych sytuacji w których wszystkie wyniki są jednakowo ‚możliwe’ to nie można jest stosować do pojedynczego rzutu lub wielokrotnych rzutów asymetryczną monetą” [1].

2.

„Klasyczna definicja wydaje się być cyrkularną, ponieważ odnosi się ona do ‚jednakowo możliwych przypadków” – i tak prawdopodobieństwo definiowane jest z użyciem pojęcia prawdopodobieństwa” [1].

3.

„Według frekwentystów skończonego ciągu zdarzeń, moneta, która nigdy nie była rzucona i stąd brak dla niej jakiegokolwiek ewidencji zdarzeń, nie posiada w ogóle prawdopodobieństwa reszki; jednak nie można powiedzieć, że moneta, która nigdy nie była pomierzona nie posiada z tego powodu średnicy” [7].

4.

„Według teorii częstościowej nie można stosować pojęcia prawdopodobieństwa do pojedynczych zdarzeń” [1], takich jak pojedynczy rzut monetą. „... moneta rzucona jeden raz dostarcza względną częstość reszki albo 0 albo 1, niezależnie od tego jak silnie jest asymetryczna, ... , jest to tzw. ‚problem pojedynczego przypadku’ [7]. Czy mamy z faktu, że w pojedynczym rzucie wypadł orzeł wnioskować, iż prawdopodobieństwo orła wynosi 1? Taki właśnie wniosek sugeruje teoria częstościowa. Byłby on wysoce pochopny i przedwczesny. Rozpatrzmy, jako drugi przykład, przykład lekarza, który chce określić prawdopodobieństwo zachorowania na nowotwór płuc pod wpływem palenia papierosów. Lekarz ten zaczyna w swojej bazie danych gromadzić informacje o pacjentach. Załóżmy, że na początku zetknął się tylko z jednym pacjentem, który mimo palenia papierosów jest zdrowy. Wniosek, jaki sugeruje częstościowa interpretacja miałby w takim przypadku postać:

„prawdopodobieństwo nowotworu płuc u osoby palącej wynosi zero”. Takie wnioskowanie byłoby oczywiście nonsensem. Powyższe przykłady pokazały, że częstościowej interpretacji prawdopodobieństwa rzeczywiście nie można stosować do pojedynczego przypadku. Co jednak, jeśli przypadków jest więcej, 2, 3, ..., 10? Okazuje się, że częstościowa interpretacja prawdopodobieństwa również dla małej liczby przypadków daje wątpliwe, a czasem zupełnie niewiarygodne wyniki. Załóżmy, że rzuciliśmy monetą 5 razy i za każdym razem wypadł orzeł. Interpretacja częstościowa sugeruje w takim przypadku prawdopodobieństwo orła równe 1 a reszki równe 0. Załóżmy, że lekarz z podanego wcześniej przykładu ma swojej bazie danych 10 pacjentów, którzy mimo palenia znacznej liczby papierosów dziennie nie mają nowotworu płuc. Częstościowa interpretacja sugeruje w takim przypadku prawdopodobieństwo równe 1 dla hipotezy: „Palenie papierosów nie wywołuje nowotworu płuc”. Oczywiście, nikt z nas nie zaakceptowałby takiego wyniku badawczego. Widać więc wyraźnie, że w przypadku małej liczby dowodów interpretacja częstościowa nie może być stosowana. Wiarygodność dostarczanych przez nią wyników jest wątpliwa. Stąd wielu naukowców uważa, że „prawdopodobieństwa pojedynczych przypadków są nonsensami”, Hajek w [7]. Jednak nie jest to prawdą. Zaproponowana w niniejszym artykule nowa interpretacja prawdopodobieństwa pokaże, że problem pojedynczego przypadku może być wiarygodnie rozwiązany.

## 5. Osobliwości częstościowej interpretacji prawdopodobieństwa przy jednorodnych dowodach.

Rozważmy ponownie przykład lekarza, który zamierza określić prawdopodobieństwo hipotezy  $h$  „palenie papierosów zwiększa ryzyko nowotworu płuc”. Anty-hipotezą  $NOT\ h = \bar{h}$  jest „palenie papierosów nie zwiększa ryzyka nowotworu płuc”. Lekarz po pewnym okresie zbierania danych posiada w swej bazie dane o  $n = 50$  pacjentach, z których wszyscy palili papierosy ale żaden nie miał nowotworu płuc. Z tego względu wszyscy ci pacjenci są potwierdzeniami anty-hipotezy  $\bar{h}$ , stąd  $n_h = 0$  a  $n_{\bar{h}} = 50$ . Odpowiednio, stosując częstościową interpretację, otrzymujemy następujące wartości prawdopodobieństw:

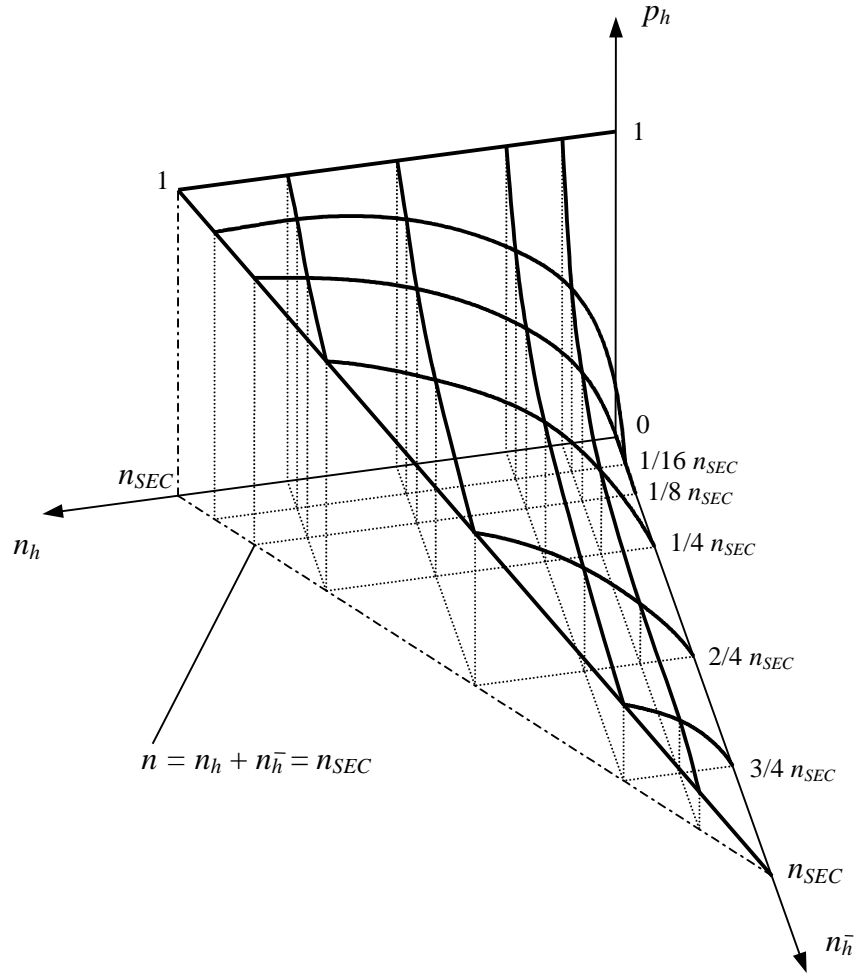
$$p_h = n_h/n = 0/50 = 0 \quad \text{oraz} \quad p_{\bar{h}} = n_{\bar{h}}/n = 50/50 = 1.$$

Z powyższego wynika trudny do zaakceptowania wniosek: „palenie papierosów nie zwiększa ryzyka nowotworu płuc” z prawdopodobieństwem 1. Zgodzimy się chyba wszyscy ze stwierdzeniem, że powyższy wniosek byłby fałszywy. Jednak taki właśnie bezpośredni wniosek sugeruje nam częstościowa interpretacja prawdopodobieństwa. Jak zobaczymy później, nowa, kompletnościowa interpretacja sugeruje wniosek znacznie rozsądniejszy. Podobną sytuację można zaobserwować na przykładzie rzutu monetą. Załóżmy, że we wszystkich  $n = 10$  rzutach moneta, które wykonaliśmy wypadła reszka. Bezpośredni wniosek, jaki sugeruje nam interpretacja częstościowa dotyczący prawdopodobieństwa reszki  $p_h$  to:

$$p_h = n_h/n = 10/10 = 1.$$

Wniosek ten oznaczałby całkowitą dominację reszki w monecie, co praktycznie się nie zdarza. Co jest przyczyną takich błędnych sugestii dostarczanych przez interpretację częstościową? Przyczyną jest błędny wzór  $p_h = n_h/n$  do obliczania prawdopodobieństwa podawany przez tę interpretację. Nielogiczność tego wzoru ilustruje rys. 1.

Wzór  $p_h = n_h/n$  pozornie wydaje się być wzorem liniowym względem  $n_h$ , bowiem w jego liczniku występuje liczba dowodów popierających hipotezę  $h$ . Stąd można by sądzić, że każde kolejne potwierdzenie tej hipotezy zwiększa jej prawdopodobieństwo o tą samą



$$p_h = \frac{n_h}{n} = \frac{n_h}{n_h + n_{\bar{h}}}$$

Rysunek 1: Funkcyjna powierzchnia prawdopodobieństwa hipotezy  $h$  obliczanej na podstawie wzoru  $p_h = n_h/n = n_h/(n_h + n_{\bar{h}})$  sugerowanego przez częstościową interpretację prawdopodobieństwa. Oznaczenia:  $n$  – sumaryczna liczba potwierżeń (dowodów),  $n_h$  – liczba potwierżeń hipotezy  $h$ ,  $n_{\bar{h}}$  – liczba potwierżeń anty-hipotezy  $\bar{h}$ ,  $n_{SEC}$  – pewna stała liczba potwierżeń ( $n_{SEC} = \text{const}$ ).

wartość  $1/n$ . Jednak tak nie jest. Liczba potwierżeń  $n_h$  hipotezy występuje także w mianowniku, mamy więc:  $p_h = n_h/(n_h + n_{\bar{h}})$ . Z tego względu zależność prawdopodobieństwa hipotezy  $p_h$  od liczby  $n_h$  jej potwierżeń jest nieliniowa, a to właśnie wywołuje szereg nielogiczności. Zwróćmy np. uwagę na sytuację, gdy brak jest jakichkolwiek dowodów potwierdzających prawdziwość anty-hipotezy  $\bar{h}$  (to znaczy  $n_{\bar{h}} = 0$ ), ale istnieją dowody poparcia hipotezy  $h$  ( $n_h > 0$ ). Sytuacji tej odpowiada pionowa ściana na rys. 1. Jeżeli nie posiadamy żadnych dowodów ani na poparcie hipotezy  $h$  ani na poparcie jej anty-hipotezy  $\bar{h}$ , to znaczy  $n_h = 0$  oraz  $n_{\bar{h}} = 0$ , wówczas według częstościowej interpretacji prawdopodobieństwo  $p_h$  hipotezy  $h$  jest nieokreślone. Jeśli jednak zdobędziemy chociaż jeden, jedyny dowód popierający tę hipotezę to wartość jej prawdopodobieństwa staje się znana i ma ono wartość  $p_h = 1$ , co przecież oznacza pewność. Jeżeli zdobędziemy np. 10 dowodów na poparcie hipotezy  $h$  przy braku dowodów na poparcie jej anty-hipotezy  $\bar{h}$  ( $n_h = 10$ ,  $n_{\bar{h}} = 0$ ) to prawdopodobieństwo hipotezy nie ulega żadnym zmianom i dalej

wynosi  $p_h = 1$ . Jeżeli posiadać będziemy 1000 dowodów na poparcie hipotezy  $h$ , przy braku dowodów dla anty-hipotezy  $\bar{h}$  ( $n_h = 1000$ ,  $n_{\bar{h}} = 0$ ), to wartość  $p_h$  pozostaje w dalszym ciągu równa 1. Zastanówmy się jednak, czy liczba dowodów na poparcie dowolnej hipotezy powinna wpływać na wartość prawdopodobieństwa  $p_h$  tej hipotezy czy nie? Według częstościowej interpretacji prawdopodobieństwa, w pewnych sytuacjach liczba dowodów jest bez znaczenia, w innych natomiast ma znaczenie (gdy początkowa liczba dowodów  $n_h > 0$  oraz  $n_{\bar{h}} > 0$ ). Czy jest to logiczne? Na pewno nie. Jak pokazane będzie w następnych rozdziałach, nowa kompletnościowa interpretacja prawdopodobieństwa daje pod tym względem o wiele bardziej logiczne, zrozumiałe i akceptowalne wyniki.

## 6. Fluktuacje prawdopodobieństwa przy małej liczbie dowodów.

Autor wykonał eksperyment dziesięciu rzutów monetą uzyskując następujące wyniki (O-orzeł, R- reszka): (O,R,O,O,O,R,O,R,R,O). W tabeli 1 pokazano wartości prawdopodobieństwa częstościowego:  $n_h/n$ .

| Wynik rzutu | $n_h$ | $n$ | $p_h = n_h/n$  |
|-------------|-------|-----|----------------|
| O           | 0     | 1   | 0              |
| R           | 1     | 2   | $1/2 = 0.500$  |
| O           | 1     | 3   | $1/3 = 0.333$  |
| O           | 1     | 4   | $1/4 = 0.250$  |
| O           | 1     | 5   | $1/5 = 0.200$  |
| R           | 2     | 6   | $2/6 = 0.333$  |
| O           | 2     | 7   | $2/7 = 0.286$  |
| R           | 3     | 8   | $3/8 = 0.375$  |
| R           | 4     | 9   | $4/9 = 0.444$  |
| O           | 4     | 10  | $4/10 = 0.400$ |

Tabela 1: Wartości prawdopodobieństwa  $p_h = n_h/n$  obliczanego według częstościowej interpretacji prawdopodobieństwa po każdym z kolejnych rzutów monetą z serii 10 rzutów dla serii (O,R,O,O,O,R,O,R,R,O), gdzie: R – reszka, O – orzeł,  $n_h$  – liczba reszek,  $n$  – liczba rzutów.

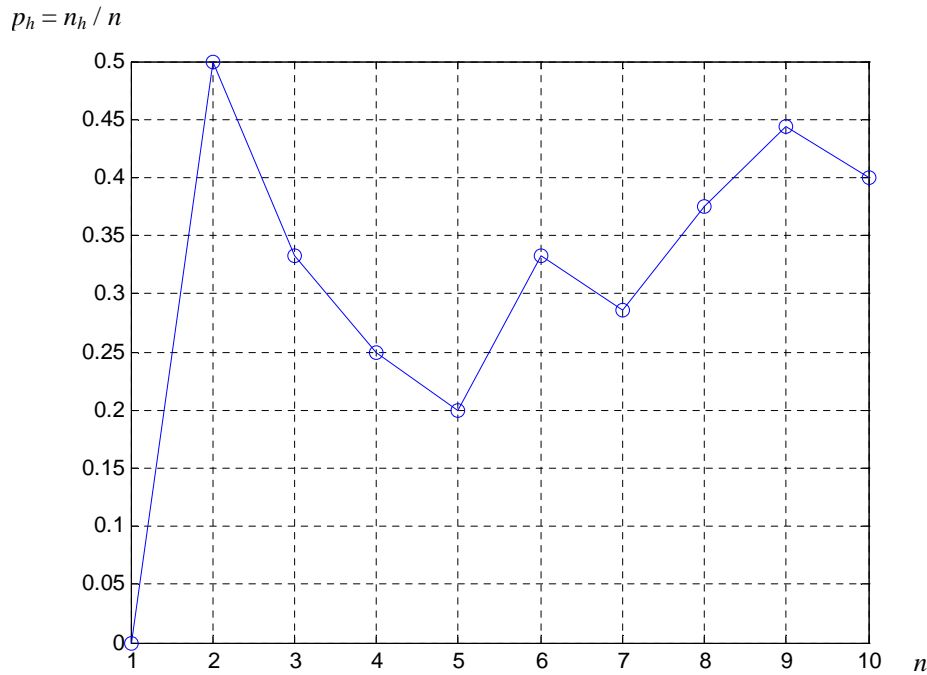
Na rys. 2 pokazano natomiast wykres wartości estymat prawdopodobieństwa  $n_h/n$  hipotezy  $h$  (w monecie dominuje reszka) podanych w tabeli 1. Wynik pojedynczego rzutu może być nazywany potwierdzeniem (pojedynczym dowodem) hipotezy  $h$  ('przeważa reszka') albo potwierdzeniem anty-hipotezy ('nie przeważa reszka' lub równoważnie 'przeważa orzeł').

Jak można zauważyć na rys. 2 wartości estymat prawdopodobieństwa ulegają znacznym wahaniom po każdym kolejnym rzucie. Estymata prawdopodobieństwa dostarczana przez interpretację częstościową przypomina więc niezdecydowaną osobę, która zbyt szybko i zbyt pochopnie wyciąga wnioski po uzyskaniu każdej nowej wiadomości. Z tego względu musi je potem znacznie korygować.

## 7. Fluktuacje prawdopodobieństwa przy dużej liczbie dowodów.

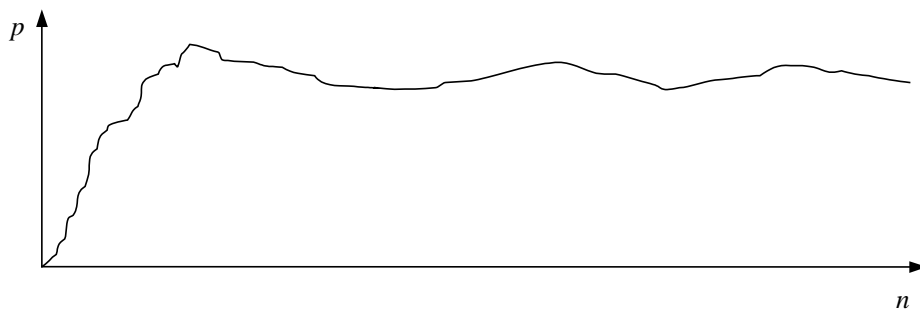
Według częstościowej interpretacji prawdopodobieństwa opierającej się na wielkiej liczbie dowodów (wyników eksperymentów, próbek) prawdziwa, dokładna wartość prawdopodobieństwa może być poznana po dokonaniu ogromnej liczby eksperymentów (teoretycznie





Rysunek 2: Wykres wartości estymat prawdopodobieństwa  $p_h = n_h/n$  hipotezy  $h$  (dominuje reszka) obliczonych po każdym kolejnym rzucie dla sekwencji (O,R,O,O,O,R,O,R,R,O) jako przykład znacznych fluktuacji prawdopodobieństwa obliczanego na podstawie interpretacji częstościowej, gdzie  $n_h$  oznacza liczbę reszek po  $n$  rzutach.

nieskończenie wielkiej liczby eksperymentów). Niestety, doświadczenia przeprowadzone przez naukowców wykazały, że nawet przy ogromnej liczbie eksperymentów nie zawsze dochodzi do ustabilizowania się prawdopodobieństwa i obserwuje się jego nieustanne fluktuacje, patrz np. Larose [14]. Zjawisko to ilustruje rys. 3.



Rysunek 3: Zjawisko fluktuacji wartości prawdopodobieństwa obliczanego na podstawie częstościowej interpretacji prawdopodobieństwa ( $p_h = n_h/n$ ) obserwowane po wielkiej liczbie prób.

Jak wykazały analizy przeprowadzone przez autora główną przyczyną znacznych fluktuacji prawdopodobieństwa pojmowanego jako częstość względna jest sama postać wzoru  $p_h = n_h / (n_h + n_{\bar{h}})$  stosowanego przez częstościową interpretację prawdopodobieństwa do obliczania prawdopodobieństwa a nie przyczyny obiektywne. Dotyczy to zarówno przypadku małej jak i wielkiej liczby eksperymentów. W przypadku nowej interpretacji prawdopodobieństwa fluktuacje przy małej liczbie dowodów są znacznie mniejsze niż w przypadku

interpretacji częstościowej a przy wielkiej liczbie praktycznie nie występują, co potwierdza występowanie zjawiska stabilizacji prawdopodobieństwa.

**8.** ‘Niesprawiedliwość’ i nielogiczność częstościowej interpretacji prawdopodobieństwa. Załóżmy, jak w punkcie 6, że dysponujemy wynikami eksperymentu rzutu monetą w formie sekwencji wyników (O,R,O,O,O,R,O,R,R,O). W tabeli 2 przedstawiono wyniki obliczeń estymaty prawdopodobieństwa  $p_h$  hipotezy  $h$  (przeważa reszka) oraz zmian  $\Delta p_h$  tego prawdopodobieństwa po każdym kolejnym rzucie, a więc po dostarczeniu kolejnych potwierdzeń bądź zaprzeczeń hipotezy.

| Wynik rzutu | $n_h$ | $n$ | $p_h = n_h/n$  | zmiana $\Delta p_h$ |
|-------------|-------|-----|----------------|---------------------|
| O           | 0     | 1   | 0              | —                   |
| R           | 1     | 2   | $1/2 = 0.500$  | +0.500              |
| O           | 1     | 3   | $1/3 = 0.333$  | -0.167              |
| O           | 1     | 4   | $1/4 = 0.250$  | -0.083              |
| O           | 1     | 5   | $1/5 = 0.200$  | -0.050              |
| R           | 2     | 6   | $2/6 = 0.333$  | +0.133              |
| O           | 2     | 7   | $2/7 = 0.286$  | -0.047              |
| R           | 3     | 8   | $3/8 = 0.375$  | +0.089              |
| R           | 4     | 9   | $4/9 = 0.444$  | +0.069              |
| O           | 4     | 10  | $4/10 = 0.400$ | -0.044              |

Tabela 2: Wartości prawdopodobieństwa  $p_h = n_h/n$  hipotezy  $h$  o dominacji reszki i zmian prawdopodobieństw  $\Delta p_h$  następujących po każdym kolejnym rzucie (względem poprzedniego rzutu) w serii 10 rzutów (O,R,O,O,O,R,O,R,R,O) obliczanych na podstawie częstościowej interpretacji prawdopodobieństwa, gdzie  $n_h$  oznacza liczbę reszek po  $n$  rzutach, R oznacza reszkę, O oznacza orła.

Jak można zauważyć w tabeli 2, w drugim rzucie wypadła reszka powodując przyrost  $\Delta p_h = 0.5$ . W dalszym ciągu reszka ta będzie nazywana reszką R2. Kolejna reszka wypadła w rzucie 6 (reszka R6) i spowodowała wzrost prawdopodobieństwa reszki  $\Delta p_h = 0.133$ , a więc o wartość znacznie mniejszą niż reszka R2. Trzecia reszka R8 spowodowała zmianę częstościowego prawdopodobieństwa o jeszcze mniejszą wartość  $\Delta p_h = 0.089$ , a ostatnia reszka R9 spowodowała zmianę najmniejszą, wynoszącą jedynie  $\Delta p_h = 0.044$ . Ponieważ każda z kolejnych reszek wywoływała różną zmianę prawdopodobieństwa  $\Delta p_h$ , to z faktu tego wynika, że każda z tych reszek ma inne znaczenie dowodowe dla oceny prawdopodobieństwa. Czy jest to logiczne i ‘sprawiedliwe’? Dlaczego wynik jednego rzutu monetą ma być ważniejszy od wyniku innego rzutu monetą? Jest to z pewnością nielogiczne i trudno byłoby zjawisko to uzasadnić. Zjawisko różnej wagi dowodów zależnie od ich pozycji w sekwencji dowodów znane jest pod nazwą ‘problemu porządku sekwencji’ (sequence-ordering problem) i zauważone zostało i opisane przez wielu naukowców, np. w [1, 7]. Jak pokazane będzie w dalszej części wykładu zjawisko to nie występuje w przypadku nowej, proponowanej interpretacji prawdopodobieństwa, gdzie każdy kolejny, pojedynczy dowód zmienia estymatę prawdopodobieństwa o taką samą wartość.

## 4 Nowa, kompletnościowa interpretacja prawdopodobieństwa

W dalszym ciągu przedstawiona zostanie nowa (według wiedzy autora) kompletnościowa interpretacja prawdopodobieństwa [16], która nie posiada szeregu wad interpretacji częstościowej. Należy przede wszystkim zwrócić uwagę na fakt, że wzór  $p_h = n_h/n$  proponowany przez częstościową interpretację prawdopodobieństwa jest częściowo błędny, choć nie całkowicie. Pozwala on obliczyć wiarygodne wartości prawdopodobieństwa tylko dla dużej lub bardzo dużej liczby próbek. Wyjaśnia to praktyczną użyteczność częstościowej interpretacji w statystyce, gdzie możemy dysponować większym materiałem ewidencyjnym. Natomiast, jak pokazano w poprzednim rozdziale wzór ten nie nadaje się do estymowania prawdopodobieństwa dla braku oraz małej liczby próbek (dowodów), a więc nie potrafi modelować pewnej klasy problemów realnie występujących w praktyce. Dodatkowo, jego nieliniowa forma matematyczna jest powodem znacznej fluktuacji prawdopodobieństwa występującej przy każdej liczbie próbek. Generalnie, częstościowa interpretacja sprawdza się w jednym zakresie liczby próbek, w innym natomiast nie. Jest więc interpretacją niedoskonałą, wymagającą korekty. Wydaje się, że przyczyną częściowej wadliwości tej interpretacji jest brak pewnego ważnego elementu w jej koncepcji. Autor proponuje nadać temu elementowi nazwę kompletnego zbioru ewidencyjnego lub w skrócie kompletności ewidencyjnej (evidential completeness – EC).

Prawdopodobieństwo ma w języku polskim dosłowny sens ‘podobieństwa do prawdy’. Stąd, jeśli chcemy określić prawdopodobieństwo rozpatrywanej hipotezy  $h$  dotyczącej jakiegoś zdarzenia na podstawie materiału ewidencyjnego potwierdzającego (bądź nie) prawdziwość tej hipotezy musimy wiedzieć jaki powinien być kompletny materiał dowodowy  $EC_h$  (kompletność ewidencyjna hipotezy), który w pełni potwierdziłby prawdziwość tej hipotezy.

Rozpatrzmy teraz przypadek dyskretny. Załóżmy, że dyskretna zmienna probabilistyczna  $X$  może w przyszłym zdarzeniu przyjąć  $k$  wartości. Na przykład, jeśli zmienna  $X$  reprezentuje wynik zdarzenia ‘rzut kostką’ to możliwych jest 6 wyników. Przed rzutem możemy zatem sformułować 6 hipotez dotyczących realizacji przyszłego zdarzenia:  $h_1^* = 1$  oczko,  $h_2^* = 2$  oczka,  $\dots$ ,  $h_6^* = 6$  oczek. Natomiast w ogólnym przypadku dyskretnym możemy sformułować  $k$  hipotez dotyczących wyniku zdarzenia. Zbiór  $H$  wszystkich hipotez będzie mieć formę jak poniżej.

$$H = \{h_1^*, h_2^*, \dots, h_k^*\}$$

Każdy zbiór hipotez o skończonej liczności można przetransformować w 2-elementowy, binominalny zbiór hipotez:

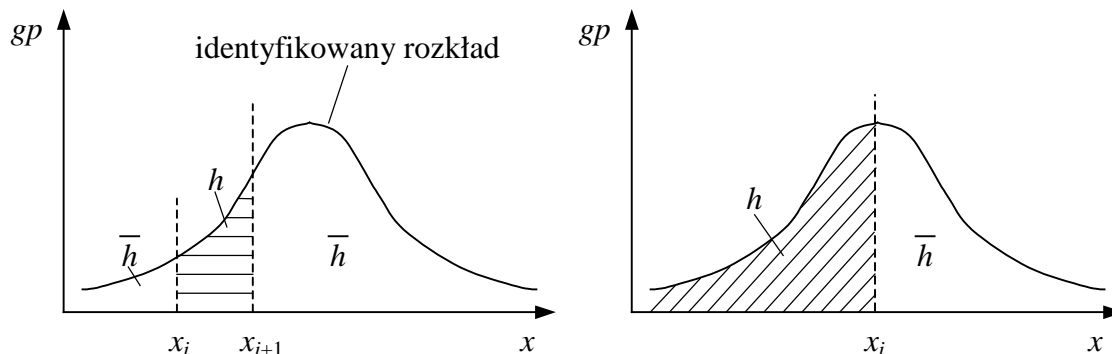
$$H = \{h, NOT h\} = \{h, \bar{h}\}$$

gdzie:  $h = h_i^*$ ,  $\bar{h} = NOT h_i^*$ ,  $i = 1, \dots, k$ .

Na przykład, w przypadku rzutu kostką możemy oddzielnie badać prawdopodobieństwo hipotezy  $h_1^* = 1$  oczko oraz anty-hipotezy  $NOT h_1^*$  (NIE 1 oczko). Następnie, sukcesywnie możemy badać prawdopodobieństwo hipotezy  $h_2^* = 2$  oczka oraz anty-hipotezy  $NOT h_2^*$  (NIE 2 oczka), etc. W ten sposób problem z większą liczbą hipotez niż 2 może być przetransformowany w  $k - 1$  sukcesywnie, kolejno rozwiązywanych pod-problemów binomialnych typu  $H = \{h, NOT h\}$ , a więc problemów typu rzut monetą, gdzie wynikiem rzutu może być albo reszka  $h$  albo orzeł  $NOT h$ . Jeżeli zmienna probabilistyczna  $X$  jest zmienną ciągłą to można dokonać jej dyskretyzacji na  $k$  pod-przedziałów i zamienić

zadanie identyfikacji rozkładu gęstości prawdopodobieństwa tej zmiennej na  $k - 1$  sukcesywnie rozwiązywanych zadaniach identyfikacji prawdopodobieństwa typu binominalnego za zbiorem hipotez  $H = \{h, NOT h\}$ , rys. 4.

Hipoteza  $h$ :  $x \in [x_i, x_{i+1}]$



Rysunek 4: Transformacja zadania określenia rozkładu gęstości prawdopodobieństwa ciągłej zmiennej probabilistycznej  $X$  na  $k - 1$  zadań z binarnym zbiorem hipotez  $H = \{h, NOT h\}$  typu rzut monetą. Oznaczenia:  $gp$  – gęstość prawdopodobieństwa,  $\bar{h} = NOT h$  – anty-hipoteza.

Zmniejszając granulację pod-przedziałów można zwiększyć dokładność przybliżenia zmiennej ciągłej przez zmienną dyskretyzowaną. Widać więc, że binomialny problem zdarzenia o 2 możliwych wynikach typu rzut monetą jest podstawowym problemem probabilistycznym, a jego rozwiązanie jest kluczem do rozwiązania innych, bardziej złożonych problemów z większą niż 2 liczbą hipotez. Z tego względu nowa interpretacja prawdopodobieństwa przedstawiona zostanie na przykładzie problemu binominalnego. Będzie ona w wykładzie tłumaczona, w celu ułatwienia jej zrozumienia, na przykładzie rzucania monetą. **Nie należy jednak wyciągać z tego mylnego wniosku, że celem wykładu jest analiza problemu rzucania monety. W wykładzie rozważany jest ogólny problem binomialny z dwoma hipotezami. Przykład rzucania moneta jest jedynie jedną z możliwych ilustracji tego problemu.**

Inne przykłady problemów binominalnych:

- hipoteza – polska reprezentacja olimpijska zdobędzie podczas następnej olimpiady co najmniej 2 złote medale,  
antyhipoteza – polska reprezentacja zdobędzie mniej niż 2 złote medale,
- hipoteza – w przyszłym roku średnia cena paliwa V-Power Diesel przekroczy 5 zł./litr,  
antyhipoteza – w przyszłym roku średnia cena paliwa V-Power Diesel nie przekroczy 5 zł./litr,
- hipoteza – spożywanie znacznych ilości białka znacznie zwiększa ryzyko choroby Alzheimera,  
antyhipoteza – spożywanie znacznych ilości białka nie zwiększa znacznie ryzyka choroby Alzheimera.

Według autora, **prawdopodobieństwo raczej nie powinno być przypisywane zdarzeniom** lub stanom **lecz hipotezom dotyczącym możliwych wyników tych zdarzeń** lub możliwym formom stanów. Na przykład, przed rzuceniem monety możemy

sformułować 2 hipotezy dotyczące wyniku eksperymentu: hipotezę  $h$  (dominacja reszki w monecie, lub w skrócie: reszka) oraz anty-hipotezę  $NOT h$  o dominacji NIE-reszki czyli orła. Natomiast po zajściu zdarzenia możemy mówić jedynie o jego realizacji  $r$ : realizacją może być wypadnięcie albo reszki ( $r = \text{reszka}$ ) albo orła ( $r = \text{orzeł}$ ).

Realizacja może być tylko jedna a prawdopodobieństwo realizacji  $p_r$  może przyjmować albo wartość 1 (pewność) albo 0. Nie może natomiast przyjmować wartości ułamkowych tak typowych dla szacunku prawdopodobieństwa. Realizacja pojedynczego rzutu dostarcza pojedynczego potwierdzenia (pojedynczego dowodu) albo dla hipotezy  $h$  (przeważa reszka) albo dla anty-hipotezy  $NOT h$  (przeważa orzeł). Jeżeli zrealizowaliśmy  $n$  rzutów monetą, w których w  $k$  rzutach wypadła reszka a w  $(n-k)$  orzeł, to możemy powiedzieć, że dysponujemy  $k$  potwierdzeniami dla hipotezy  $h$  (przeważa reszka) oraz  $(n-k)$  potwierdzeniami anty-hipotezy  $NOT h$  (przeważa orzeł). Powstaje pytanie, jak mocny jest posiadany przez nas zbiór ewidencyjny  $E_h$  wspierający hipotezę oraz zbiór  $E_{NOT h}$  wspierający anty-hipotezę i czy na podstawie posiadanej liczby  $n$  dowodów (np.  $n = 1$ ,  $n = 5$ ,  $n = 17$ , etc.) wolno nam wnioskować o prawdopodobieństwie hipotezy  $h$  (przeważa reszka)? A jeśli tak, to czy prawdopodobieństwo jednej i tej samej hipotezy określone na podstawie różnej liczby dowodów (wyników eksperymentów) będzie jednakowo wiarygodne? Potrzebna jest więc odpowiedź na pytanie o konieczną liczbę  $n$  zbioru ewidencyjnego. Jak już wcześniej podano, zbiór  $EC_h$  dowodów w pełni potwierdzający prawdziwość hipotezy  $h$  i wykluczający minimalne nawet prawdopodobieństwo anty-hipotezy  $NOT h$  ( $p_h = 1$ ,  $p_{NOT h} = 0$ ) nazywany będzie kompletnym zbiorem ewidencyjnym hipotezy, w skrócie kompletnością ewidencyjną. Zbiór ten może mieć w niektórych problemach formę idealną. Rozpatrzmy dla przykładu kryminalną sprawę morderstwa.

Policja wytypowała kilka osób podejrzanych o jego popełnienie. Jedną z tych osób jest osoba A. Binomialny zbiór hipotez ma w tym wypadku postać  $H = \{h_A, NOT h_A\}$ , gdzie  $h_A$  oznacza hipotezę ‘osoba A popełniła morderstwo’ a anty-hipoteza  $NOT h_A$  ma sens ‘osoba A nie popełniła morderstwa’ lub też ‘osoby inne niż osoba A popełniły morderstwo’. Kompletność ewidencyjna  $EC_h$  oznacza w tym wypadku taki zbiór dowodów względem osoby podejrzanej (dowolnej osoby podejrzanej, niekoniecznie osoby A) o popełnienie morderstwa, który w pełni wykazywałby (dowodziłby), że osoba podejrzana, niezależnie od tego, kto jest tą osobą, popełniła morderstwo. Poniżej podany jest zbiór takich dowodów, który według autora (nie będącego ekspertem sądowym) wydaje się być zbiorem kompletnym i w pełni wykazującym winę podejrzanego.

$EC_h = \{$  *osoba podejrzana widziana była przez kilku świadków podczas popełniania morderstwa, osoba podejrzana rozpoznana została przez wszystkich tych świadków podczas eksperymentu identyfikacyjnego, osoba podejrzana nie ma alibi na czas morderstwa, osoba podejrzana miała istotne motywy do popełnienia morderstwa, w miejscu popełnienia morderstwa wykryto materiał genetyczny osoby podejrzanej, na narzędziu popełnienia morderstwa wykryto materiał genetyczny osoby podejrzanej*  $\}$

Jeżeli w przypadku osoby podejrzanej A dysponowalibyśmy wszystkimi dowodami zawartymi w kompletności ewidencyjnej  $EC_h$  to prawdopodobieństwo hipotezy  $h$  ‘A jest mordercą’ byłoby całkowite i równe 1. Jeżeli jednak dysponowalibyśmy niepełnym zbiorem dowodów  $E_{h_A}$  (evidential set) jak np. poniżej,

$E_{h_A} = \{$  *osoba A miała istotne motywy do popełnienia morderstwa, osoba A nie ma alibi na czas popełnienia morderstwa*  $\},$

to prawdopodobieństwo hipotezy  $h$  ‘osoba A jest mordercą’ byłoby tylko ułamkowe. Należy zauważyć, że poszczególne dowody w zbiorze ewidencyjnym  $E_{h_A}$  są obiektywnymi, pewnymi faktami stwierdzonymi przez policję (ich prawdopodobieństwo wynosi zatem 1 i nie podlega ocenie). Ocenie podlegać może jedynie stopień prawdopodobieństwa hipotezy  $h$ . Stopień prawdopodobieństwa tej hipotezy może być w przybliżeniu oceniony (i jest w praktyce oceniany) przez ekspertów sądowych i kryminalistycznych. Podobnie, waga poszczególnych dowodów także musi być oceniona (i jest w praktyce oceniana) przez ekspertów. **W ogólnym przypadku waga (znaczenie) poszczególnych dowodów może być zróżnicowane.** Niekiedy, jak w przypadku rzutu monetą, wynik każdego rzutu zazwyczaj uznawany jest jako jednakowo ważny.

Istnieje wiele problemów binomialnych ze skończonym zbiorem ewidencyjnym. Rozpatrzmy następujący problem:

hipoteza: pierwszą osobą jaką spotkam na korytarzu mego wydziału będzie kobieta,  
 antyhipoteza: pierwszą osobą jaką spotkam na korytarzu mego wydziału będzie mężczyzna.

W tym wypadku kompletny zbiór ewidencyjny zawiera skończoną liczbę osób studiujących i pracujących na moim wydziale.

Zajmijmy się jednak teraz ponownie problemem, jaki powinien być kompletny zbiór ewidencyjny w przypadku problemu binominalnego  $H = \{h_A, NOT h_A\}$  typu rzut monetą? W powyższym przypadku idealny, kompletny zbiór ewidencyjny  $EC_h$  musiałby składać się z wyników nieskończenie wielkiej liczby rzutów monetą (dowodów). Jest oczywistym, że takiej liczby dowodów nie można zebrać. Podobna sytuacja istnieje także w innych problemach probabilistycznych. Z tego względu w wielu zadaniach praktycznych będziemy musieli się posługiwać nie idealną lecz przybliżoną kompletnością ewidencyjną, która zostanie nazwana **satysfakcjonującą kompletnością ewidencyjną** (satisfactory evidential completeness) i oznaczona będzie skrótowo SEC. Zbiór satysfakcjonujący jest zbiorem takich dowodów, które co prawda nie gwarantują całkowitej prawdziwości hipotezy dotyczącej rozpatrywanego zdarzenia lub stanu ale zapewniają jej prawdziwość w stopniu zadowalająco wysokim, np. w stopniu 0.99 lub 0.95, etc; mówiąc ogólnie, w stopniu satysfakcjonującym eksperta problemu. W przypadku problemu binominalnego typu rzut monetą konieczne jest więc określenie pewnej skończonej liczby eksperymentów  $n_{SEC}$ , która stanowiłaby satysfakcjonującą kompletność ewidencyjną. Liczba ta może być też rozumiana jako pewien model (reprezentacja) nieskończoności, to jest jako liczba, która w danym, specyficznym problemie satysfakcjonująco dobrze reprezentuje (zastępuje) nieskończoność. Możliwość taką daje tzw. Chernoff bound [3, 9]. Chernoff wyprowadził dla problemu binominalnego wzory, które, po przekształceniu pozwalają obliczyć minimalną, satysfakcjonującą liczbę dowodów (w przypadku rzutu monetą: minimalną satysfakcjonującą liczbę rzutów) na podstawie której można określić prawdopodobieństwo hipotezy  $h$  z pewną założoną wiarygodnością). Poniżej podano wzór Chernoffa (4).

$$n_{SEC} \geq \frac{1}{(p_{hc} - 0.5)^2} \ln \frac{1}{\sqrt{\epsilon}} \quad (4)$$

Oznaczenia:

$\epsilon$  – założona, maksymalna, probabilistyczna niepewność (błąd) wykazania hipotezy  $h$  (np. o dominacji reszki) jakiej nie chcemy przekroczyć. Jeśli przyjmiemy np.  $\epsilon = 0.01$  to dokładność wykazania hipotezy  $h$  na podstawie jednej tylko serii  $n_{SEC}$  prób będzie

wynosiła co najmniej  $1 - \epsilon = 0.99$ . Praktycznie oznacza to, że jeśli wykonalibyśmy 100 serii rzutów monetą (w każdej serii  $n_{SEC}$  rzutów) to co najwyżej tylko w jednej serii ( $\epsilon = 0.01$ ) jej wynik nie potwierdziłby prawdziwości hipotezy  $h$  a potwierdziłby antyhipotezę  $NOT\ h = \bar{h}$ . Natomiast w pozostałych 99 seriach po  $n_{SEC}$  rzutów prawdziwość hipotezy (np. o przewadze reszki) została by potwierdzona.

$p_{hc}$  – oznacza założoną przez nas wartość prawdopodobieństwa hipotezy ( $p_{hc} \geq 0.5$ ), którą chcemy sprawdzić i wykazać na podstawie  $n_{SEC}$  dowodów. Jeżeli w przypadku monety przypuszczamy (np. na podstawie wstępnych eksperymentów) że dominacja reszki, czyli asymetryczność monety na rzecz reszki wynosi około 0.55 to zakładamy  $p_{hc} = 0.55$  i przy tym założeniu obliczamy liczbę rzutów  $n_{SEC}$ , liczbę co najmniej konieczną do zadowalająco wiarygodnego wykazania prawdziwości hipotezy o dominacji reszki. Jeżeli po wykonaniu serii  $n_{SEC}$  rzutów monetą okazałoby się jednak, że wykazana prawdziwość hipotezy  $p_h$  faktycznie przewyższa założony minimalny próg  $p_{hc}$  (np.  $p_h = 0.57$  przy  $p_{hc} = 0.55$ ) to fakt ten oznaczałby, że zrealizowana liczba prób  $n_{SEC}$  była zadowalająco duża. Gdybyśmy natomiast po wykonaniu pierwotnie obliczonej liczby  $n_{SEC}$  wymaganych prób otrzymali wartość  $p_h$  niższą założonej wartości  $p_{hc}$  (np.  $p_h = 0.53$  przy  $p_{hc} = 0.55$ ) to należałoby obliczyć nową wartość  $n_{SEC}$  wymaganej liczby prób dla zaktualizowanego, odpowiednio niższego założenia  $p_{hc}$  (np. dla  $p_{hc} = 0.529$ ). Nowa wartość  $n_{SEC}$  będzie wtedy wyższa niż obliczona pierwotnie, co będzie oznaczało konieczność wykonania dodatkowych uzupełniających prób rzutów monetą. Jeżeli pozostalibyśmy jednak przy pierwotnej liczbie  $n_{SEC}$  prób to będzie oznaczało to obniżoną dokładność  $(1 - \epsilon)$  wykazania prawdziwości hipotezy.

Poniżej w tabeli 3 podano kilka przykładów obliczenia koniecznej liczby eksperymentów w problemie binominalnym typu np. rzut monetą.

|           |           |        |       |       |       |       |
|-----------|-----------|--------|-------|-------|-------|-------|
| $p_{hc}$  | 0.501     | 0.510  | 0.520 | 0.530 | 0.550 | 0.600 |
| $n_{SEC}$ | 2 302 586 | 23 026 | 5 757 | 2 559 | 921   | 231   |

Tabela 3: Przykłady wartości satysfakcjonującej kompletności dowodowej  $n_{SEC}$  w problemie binominalnym (np. wymaganej minimalnej liczby rzutów monetą) dla maksymalnego błędu  $\epsilon = 0.01$  oraz dokładności  $(1 - \epsilon) = 0.99$ .

Jak pokazują przykłady w tabeli 3, w miarę, gdy moneta staje się coraz bardziej symetryczna (prawdopodobieństwo reszki zbliża się do 0.5) liczba rzutów monetą  $n_{SEC}$  wymagana dla zadowalająco dokładnego zidentyfikowania prawdopodobieństwa reszki szybko i gwałtownie rośnie. Identyfikacja prawdopodobieństwa staje się więc coraz trudniejsza. Natomiast, im moneta jest bardziej asymetryczna (im bardziej prawdopodobieństwo reszki odbiega od wartości 0.5) tym mniej rzutów potrzebnych jest dla zadowalająco pewnego stwierdzenia przewagi reszki (lub orła). Identyfikacja prawdopodobieństwa staje się więc łatwiejsza. Przypomina to identyfikację bliźniaków. Duża różnica cech między bliźniakami ułatwia ich rozróżnienie, natomiast mała różnica bardzo utrudnia.

W dalszym ciągu przedstawione zostaną propozycje definicji prawdopodobieństwa hipotezy. Definicje te pokazują, że w dużej liczbie przypadków nie będziemy mogli określić (zidentyfikować) dokładnej wartości prawdopodobieństwa  $p_h$  hipotezy  $h$  lecz jedynie zakres  $[p_{hmin}, p_{hmax}]$  w którym wartość ta leży. Jednak, chociaż dokładnej wartości prawdopodobieństwa często nie możemy poznać, to jak zobaczymy, posiadany materiał ewidencyjny, zależnie od swojej wielkości, pozwoli nam na przybliżone określenie prawdopodobieństwa hipotezy.

**Minimalne prawdopodobieństwo**  $p_{hmin}$  **hipotezy**  $h$  dotyczącej rozpatrywanego zdarzenia lub stanu rzeczy jest stopniem zgodności (podobieństwa) posiadanego ewidencyjnego zbioru  $E_h$  dowodów potwierdzających prawdziwość tej hipotezy ze zbiorem dowodów zawartych w kompletnym zbiorze ewidencyjnym  $EC_h$  zawierającym zbiór dowodów wymaganych dla całkowitego wykazania prawdziwości hipotezy  $h$ .

**Maksymalne prawdopodobieństwo**  $p_{hmax}$  **hipotezy**  $h$  równe jest 1 minus minimalne prawdopodobieństwo  $p_{NOT hmin}$  anty-hipotezy  $NOT h$ , wzór (5).

$$p_{hmax} = 1 - p_{NOT hmin} \quad (5)$$

**Minimalne prawdopodobieństwo**  $p_{NOT hmin}$  **anty-hipotezy**  $NOT h$  jest stopniem zgodności (podobieństwa) posiadanego zbioru  $E_{NOT h}$  dowodów potwierdzających prawdziwość anty-hipotezy z kompletnym zbiorem ewidencyjnym  $EC_{NOT h}$  wymaganym dla pełnego wykazania prawdziwości anty-hipotezy  $NOT h$ .

**Maksymalne prawdopodobieństwo**  $p_{NOT hmax}$  **anty-hipotezy**  $NOT h$  równe jest 1 minus minimalne prawdopodobieństwo  $p_{hmin}$  hipotezy  $h$ , wzór (6).

$$p_{NOT hmax} = 1 - p_{hmin} \quad (6)$$

**Dokładną wartość prawdopodobieństwa**  $p_h$  **hipotezy**  $h$  (a także dokładną wartość prawdopodobieństwa  $p_{NOT h}$  anty-hipotezy  $NOT h$ ) **możemy określić** tylko wówczas, gdy spełniony będzie warunek (7), to znaczy, gdy minimalne prawdopodobieństwa hipotezy i anty-hipotezy sumują się do jedności.

$$\text{IF } (p_{hmin} + p_{NOT hmin} = 1) \text{ THEN } [(p_{hmin} = p_h) \text{ AND } (p_{NOT hmin} = p_{NOT h})] \quad (7)$$

**Dokładnej wartości prawdopodobieństwa hipotezy**  $h$  **nie możemy określić**, a więc i poznać, jeżeli minimalne prawdopodobieństwa hipotezy i anty-hipotezy nie sumują się do jedności, to znaczy gdy występuje sytuacja określona wzorem (8).

$$p_{hmin} + p_{NOT hmin} < 1 \quad (8)$$

Przyczyną takiej sytuacji jest niedostatek materiału dowodowego (ewidencyjnego). Niestety, z taką sytuacją często mamy do czynienia w problemach rzeczywistych. Ogólnie wartości prawdopodobieństw w problemie binominalnym podlegają ograniczeniom podanym poniżej. W problemach multinomialnych ograniczenia będą inne.

$$0 \leq p \leq 1 \quad (9)$$

$$0 \leq p_{hmin} + p_{hmax} \leq 2 \quad (10)$$

$$0 \leq p_{NOT hmin} + p_{NOT hmax} \leq 2 \quad (11)$$

$$0 \leq p_{hmin} + p_{NOT hmin} \leq 1 \quad (12)$$

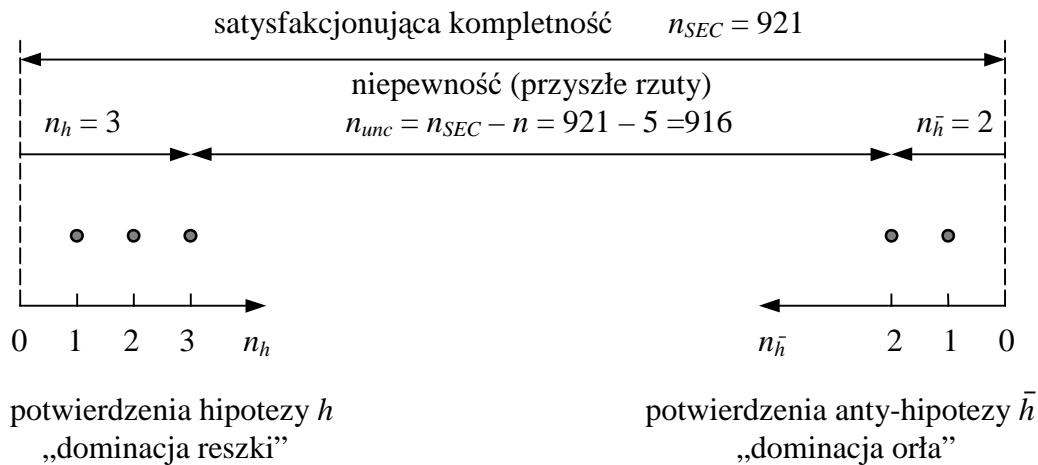
$$0 \leq p_{hmax} + p_{NOT hmax} \leq 2 \quad (13)$$

Prawdziwość powyższych ograniczeń można łatwo wykazać przy pomocy szczególnych wzorów, które zostaną podane w następnym rozdziale. Jeśli zbiór kompletności ewidencyjnej  $EC_h$  jest z pewnych powodów charakteryzujących badany problem nierealny (nigdy nie może być zebrany) to można zamiast niego użyć satysfakcjonującej kompletności ewidencyjnej SEC w celu określenia przybliżonego prawdopodobieństwa hipotezy.



## 5 Niepewność prawdopodobieństwa

Powróćmy teraz do binomialnego problemu typu rzucania monetą ze zbiorem hipotez  $H = \{h, NOT h\}$ , gdzie  $h$  oznacza hipotezę o dominacji reszki (którą można też interpretować jako hipotezę o treści ‘w następnym rzucie wypadnie reszka’). Załóżmy, że ze wstępnych eksperymentów z monetą wynika dominacja reszki i że jej przypuszczalne prawdopodobieństwo powinno być nie mniejsze niż wartość  $p_{hc} = 0.55$ . Zgodnie ze wzorem Chernoffa liczba wymaganych wyników rzutów (dowodów) do określenia prawdopodobieństwa hipotezy o dominacji reszki z dokładnością  $(1 - \epsilon) = 0.99$  czyli z maksymalnym błędem  $\epsilon = 0.01$  wynosi  $n_{SEC} = 921$ . Ta liczba wyników prób stanowi satysfakcjonująca kompletność dowodową SEC rozpatrywanego problemu. **Rozpatrzmy teraz sytuację pierwszą** i załóżmy że w pierwszym eksperymencie wykonaliśmy  $n = 5$  rzutów monetą w których uzyskaliśmy  $n_h = 3$  reszki i  $n_{NOT h} = n_{\bar{h}} = 2$  orły. Zatem zbiór ewidencyjny  $E_h$  zawiera 3 potwierdzenia (dowody) dominacji reszki a zbiór ewidencyjny  $E_{NOT h} = E_{\bar{h}}$  zawiera 2 potwierdzenia dominacji orła. Należy zauważyć, że całkowita liczba posiadanych dowodów  $n = 5$  jest znacznie mniejsza od liczby  $n_{SEC} = 921$  wymaganej przez satysfakcjonującą kompletność ewidencyjną  $E_{SEC}$ . Co w takiej sytuacji niedoboru informacyjnego możemy zrobić dla określenia prawdopodobieństwa interesującej nas hipotezy? Omawianą sytuację przedstawia rys. 5.



Rysunek 5: Ilustracja sytuacji niedoboru informacyjnego (braku 916 dowodów) i spowodowanej tym niepewności w zadaniu określenia prawdopodobieństwa hipotezy  $h$  o dominacji reszki w monecie w przypadku liczby posiadanych dowodów  $n = 5$  znacznie mniejszej od liczby dowodów  $n_{SEC} = 921$  wymaganej przez satysfakcjonujący zbiór ewidencyjny SEC.

Jak pokazuje rys. 5, aby uzyskać wymaganą sumaryczną liczbę 921 dowodów trzeba wykonać z monetą jeszcze 916 prób bowiem na razie posiadamy wyniki jedynie 5 prób. Chociaż obecne wyniki wskazują na przewagę hipotezy o dominacji reszki ( $n_h = 3$ ,  $n_{\bar{h}} = 2$ ) to następne 916 rzutów monetą może ten obraz całkowicie zmienić w nieznaną nam na razie sposób. Po wykonaniu bowiem wszystkich 921 rzutów wymaganych przez satysfakcjonujący zbiór ewidencyjny może się okazać, że to nie reszka lecz orzeł dominuje w monecie. Dlatego, na podstawie tak małej liczby 5 tylko dowodów nie wolno nam formułować żadnych kategoriycznych i pochopnych wniosków o dominacji reszki. Sytuacja jest bowiem silnie niepewna, stąd formułowane wnioski muszą być bardzo ostrożne. Mimo, że liczba dowodów  $n = 5$  jest bardzo niska, to jednak dostarcza ona nam pewnej wiedzy o prawdo-

podobieństwie i wiedzę tę można i należy wykorzystać, chociaż nie możemy sobie po niej zbyt wiele obiecywać. W krańcowym przypadku możliwym jest, że wszystkie następne 916 rzutów dałyby reszki. Wówczas reszka byłaby ostatecznie poparta  $3 + 916 = 919$  dowodami. Stąd maksymalne możliwe prawdopodobieństwo reszki jakie możemy wywnioskować z posiadanej liczby 5 dowodów wynosi  $p_{hmax} = 919/921$ . Natomiast minimalne prawdopodobieństwo reszki  $p_{hmin}$  zapewnione trzema już posiadаныmi dowodami wynosi  $3/921$ . Z kolei minimalne prawdopodobieństwo anty-hipotezy o przewadze orła wynosi  $2/921$ , bowiem posiadamy obecnie 2 potwierdzenia dominacji orła. Ponieważ możliwym jest, że wszystkie następne, brakujące rzuty w liczbie 916 dadzą w wyniku orła, to maksymalne możliwe prawdopodobieństwo orła  $p_{hmax}$  wynosi  $(2 + 916)/921 = 918/921$ . Łatwo jest zauważyć, że każde potwierdzenie orła podwyższa jego minimalne prawdopodobieństwo  $p_{hmin}$  o wartość  $1/921$  (działa na jego korzyść) a zmniejsza maksymalne możliwe prawdopodobieństwo  $p_{hmax}$  'przeciwnika' czyli reszki o taką samą wartość  $1/921$  (działa więc na niekorzyść 'przeciwnika'). Natomiast każde potwierdzenie reszki podwyższa jej minimalne prawdopodobieństwo  $p_{hmin}$  o wartość  $1/921$  i o taką samą wartość zmniejsza maksymalne prawdopodobieństwo  $p_{hmax}$  przeciwnika, czyli orła. Ponieważ najprawdopodobniej pewną część z przyszłych, brakujących 916 rzutów stanowią będą reszki oraz pewną część orły, to końcowa, dokładna wartość prawdopodobieństwa  $p_h$  reszki (w sensie dokładności wzoru Chernoffa) leżała będzie gdzieś pomiędzy minimalną  $p_{hmin}$  i maksymalnie możliwą wartością  $p_{hmax}$ .

$$p_{hmin} \leq p_h \leq p_{hmax} \quad (3/921) \leq p_h \leq (919/921)$$

Odpowiednio, prawdziwa wartość prawdopodobieństwa  $p_{\bar{h}}$  anty-hipotezy  $\bar{h}$  o dominacji orła będzie leżeć gdzieś między swoją możliwą minimalną i maksymalną wartością.

$$p_{\bar{h}min} \leq p_{\bar{h}} \leq p_{\bar{h}max} \quad (2/921) \leq p_{\bar{h}} \leq (918/921)$$

Na podstawie rozpatrywanego przykładu możemy sformułować następujące, poniżej podane wnioski dotyczące wzorów.

$$p_{hmin} = n_h/n_{SEC}, \quad p_{hmax} = 1 - p_{\bar{h}min} = 1 - n_{\bar{h}}/n_{SEC} \quad (14)$$

$$p_{\bar{h}min} = n_{\bar{h}}/n_{SEC}, \quad p_{\bar{h}max} = 1 - p_{hmin} = 1 - n_h/n_{SEC} \quad (15)$$

Dokładne prawdopodobieństwa hipotezy  $h$  i anty-hipotezy  $\bar{h}$  są od siebie zależne i spełniają warunek (16).

$$p_h + p_{\bar{h}} = 1 \quad (16)$$

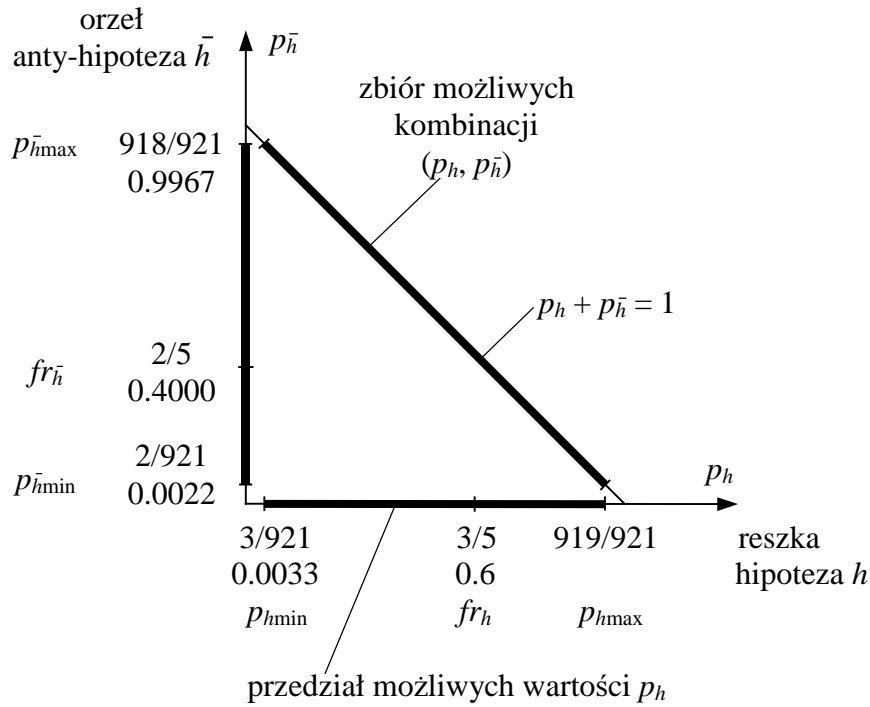
Warunek ten oznacza, że dokładne prawdopodobieństwa hipotezy i anty-hipotezy dopełniają się do jedności. Sytuację tę ilustruje rys. 6.

Jeśli do obliczenia prawdopodobieństwa hipotezy  $h$  o dominacji reszki użyjemy wzoru (3) ze strony 5, sugerowanego przez powszechnie stosowaną częstościową interpretację prawdopodobieństwa to uzyskamy wyniki jak poniżej.

$$p_h = n_h/n = 3/5, \quad p_{\bar{h}} = n_{\bar{h}}/n = 2/5 \quad (17)$$

Powstaje pytanie, dlaczego akurat wartość  $p_h = 3/5$  jest sugerowana jako estymata prawdopodobieństwa hipotezy a nie jakaś inna wartość z zakresu możliwych wartości od  $3/921$  do  $919/921$  przedstawionych na rys. 6? Czym wartość  $3/5$  wyróżnia się względem

$$n = 5, \quad n_h = 3 \text{ (reszki)}, \quad n_{\bar{h}} = 2 \text{ (NIE reszki)}$$



Wynik interpretacji kompletnościowej:  $p_h \in \left[ \frac{3}{921}, \frac{919}{921} \right]$

Wynik interpretacji częstościowej:  $p_h = \frac{n_h}{n} = \frac{3}{5}$

Rysunek 6: Ilustracja zjawiska niepewności prawdopodobieństwa  $p_h$  hipotezy oraz prawdopodobieństwa  $p_{\bar{h}}$  anty-hipotezy identyfikowanego na podstawie liczby  $n = 5$  wyników prób, która jest znacznie niższa niż liczba 921 wyników wymaganych przez satysfakcjonującą kompletność ewidencyjną  $SEC_h$ .

innych możliwych wartości? Nie ma żadnego uzasadnienia dla tak szczególnego traktowania tej wartości. Jest ona tak samo ‘dobra’ jak każda inna wartość z zakresu wartości możliwych.

**Rozpatrzmy teraz sytuację drugą**, w której dysponujemy nie małą liczbą  $n = 5$  dowodów lecz znacznie większą sumaryczną liczbą dowodów  $n = 700$ , spośród których 399 dowodów popiera hipotezę  $h$  o dominacji reszki a 301 anty-hipotezę  $\bar{h}$  o dominacji orła. Oznacza to, że minimalne prawdopodobieństwo reszki wynosi:

$$p_{hmin} = n_h/n_{SEC} = 399/921 = 0.433,$$

a minimalne prawdopodobieństwo orła wynosi:

$$p_{\bar{h}min} = n_{\bar{h}}/n_{SEC} = 301/921 = 0.430.$$

Z kolei maksymalne możliwe prawdopodobieństwo reszki wynosi:

$$p_{hmax} = 1 - p_{\bar{h}min} = 1 - (301/921) = 620/921 = 0.673,$$

a maksymalne możliwe prawdopodobieństwo orła wynosi:

$$p_{\bar{h}max} = 1 - p_{hmin} = 1 - (399/921) = 522/921 = 0.567.$$

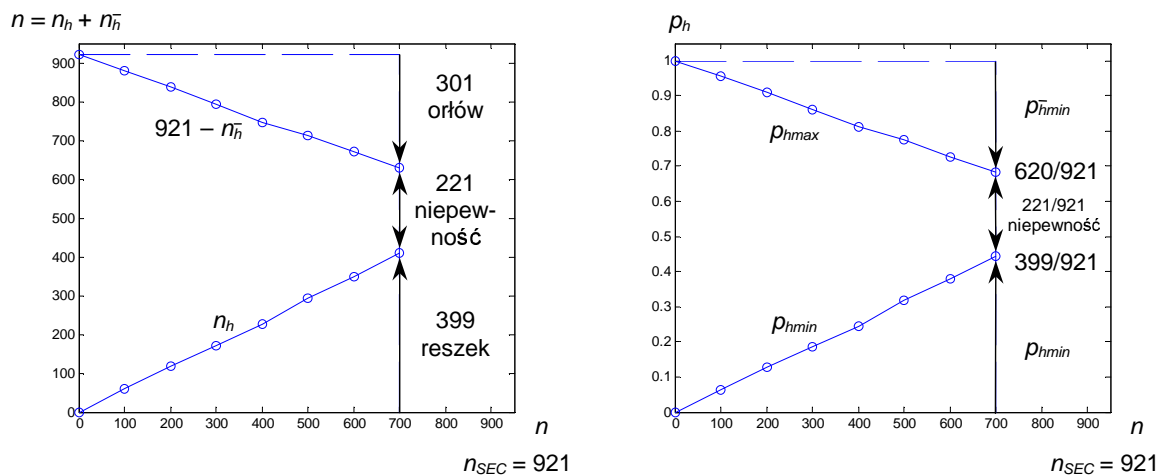
Dokładna wartość prawdopodobieństwa hipotezy  $h$  o dominacji reszki nie jest znana. Jednak wiadomo, że zawiera się ona w poznanych granicach. To samo dotyczy antyhipotezy o przewadze orła. Podsumowanie wiedzy uzyskanej na podstawie 700 rzutów monetą podane jest poniżej.

$$(399/921) \leq p_h \leq (620/921), \quad (301/921) \leq p_{\bar{h}} \leq (522/921)$$

Powyższe wyniki przedstawiono w tabeli 4, a następnie na rys. 7.

| $n$           | 0 | 100    | 200    | 300    | 400    | 500    | 600    | 700    |
|---------------|---|--------|--------|--------|--------|--------|--------|--------|
| $n_h$         | 0 | 59     | 117    | 172    | 226    | 292    | 349    | 409    |
| $n_{\bar{h}}$ | 0 | 41     | 83     | 128    | 174    | 208    | 251    | 291    |
| $p_{hmin}$    | 0 | 0.0641 | 0.1270 | 0.1868 | 0.2454 | 0.3170 | 0.3789 | 0.4441 |
| $p_{hmax}$    | 1 | 0.9555 | 0.9099 | 0.8610 | 0.8111 | 0.7742 | 0.7275 | 0.6840 |

Tabela 4: Wyniki eksperymentu rzucania monetą. Oznaczenia:  $n$  – liczba prób,  $n_h$  – liczba uzyskanych reszek,  $n_{\bar{h}}$  – liczba uzyskanych orłów,  $p_{hmin}$  – dolna granica prawdopodobieństwa  $p_h$  reszki,  $p_{hmax}$  – górna granica prawdopodobieństwa  $p_h$  reszki.



Rysunek 7: Ilustracja niepewności wyników oceny prawdopodobieństwa  $p_h$  hipotezy  $h$  o dominacji reszki uzyskanych na podstawie  $n = 700$  rzutów monetą ( $n_h = 399$  reszek i  $n_{\bar{h}} = 301$  orłów) przy wymaganej liczbie wyników  $n_{SEC} = 921$ ,  $p_{\bar{h}}$  – prawdopodobieństwo orła.

Jak pokazuje rys. 7 posiadanie wyników 700 rzutów znacznie zmniejszyło niepewność oceny prawdopodobieństwa  $p_h$  hipotezy o przewadze reszki w porównaniu z sytuacją, gdy posiadaliśmy wyniki tylko 5 rzutów, rys. 5. Niemniej niepewność ta jest dalej znaczna, bowiem 221 rzutów brakujących do pełnej liczby 921 rzutów wymaganych przez satysfakcjonującą kompletność ewidencyjną SEC może w różny sposób zmienić sytuację dowodową: może się okazać, że w monecie jednak dominuje nie reszka (obecnie  $p_{hmin} = 399/921$  przy  $p_{\bar{h}min} = 301/921$ ) lecz orzeł, o ile odpowiednia liczba przyszłych 221 rzutów wypadnie na korzyść orła, co oczywiście jest możliwe. Jeśli wyniki z rzutów monetą,  $n = 700$ ,  $n_h = 399$ ,

$n_{\bar{h}} = 301$ , użyjemy do obliczenia prawdopodobieństwa  $p_h$  na podstawie częstościowej interpretacji prawdopodobieństwa to uzyskamy wyniki przedstawione poniżej. Oznaczenie  $fr$  oznacza częstość (frequency).

$$p_h = fr_h = n_h/n = 399/700, \quad p_{\bar{h}} = fr_{\bar{h}} = 301/700$$

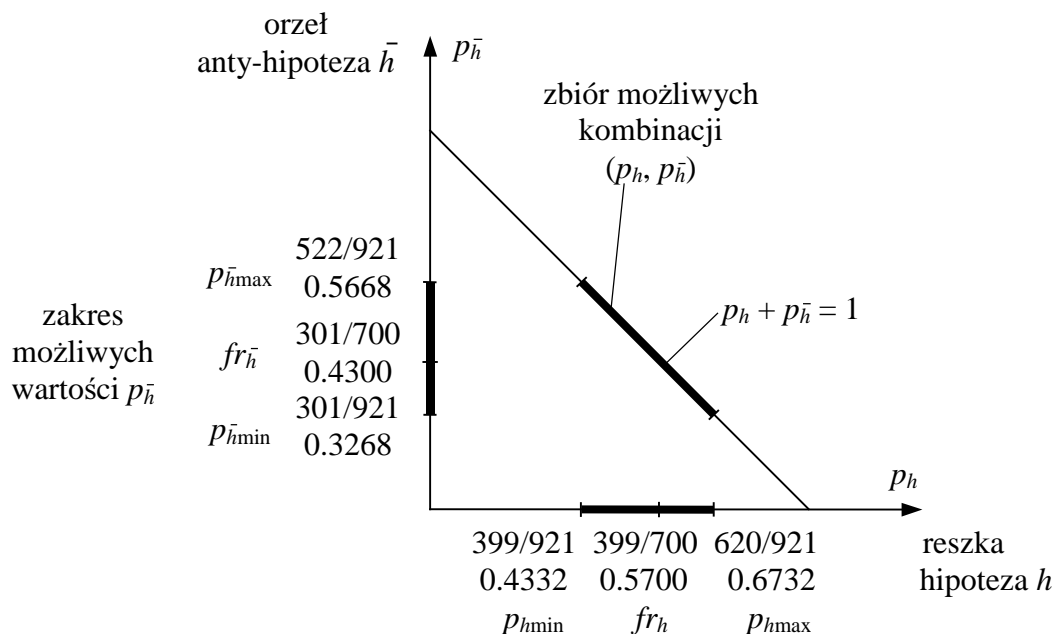
Wyniki obu interpretacji znacznie się różnią. Są one przedstawione na rys. 8.

$$n = 700, \quad n_h = 399 \text{ (reszki)}, \quad n_{\bar{h}} = 301 \text{ (orły)}$$

Wynik interpretacji kompletnościowej:

$$p_h \in \left[ \frac{n_h}{n_{SEC}}, 1 - \frac{n_{\bar{h}}}{n_{SEC}} \right] = \left[ \frac{399}{921}, 1 - \frac{301}{921} \right] = \left[ \frac{399}{921}, \frac{620}{921} \right] = [0.4332, 0.6732] \quad p_{\bar{h}} = 1 - p_h$$

$$\text{Wynik interpretacji częstościowej: } p_h = \frac{n_h}{n} = \frac{399}{700} = 0.5700 \quad p_{\bar{h}} = 1 - p_h = 0.4300$$



Rysunek 8: Porównanie wyników oceny prawdopodobieństwa  $p_h$  hipotezy  $h$  o dominacji reszki i prawdopodobieństwa  $p_{\bar{h}}$  anty-hipotezy  $\bar{h}$  o dominacji orła na podstawie wyników 700 prób (399 reszek i 301 orłów) obliczonych na podstawie interpretacji kompletnościowej oraz częstościowej.

Porównanie rys. 8 z rys. 6 pokazuje, że po zwiększeniu liczby prób rzutu monetą z 5 do 700 (przy wymaganej liczbie  $n_{SEC} = 921$ ) nastąpiło znaczne zmniejszenie się niepewności przybliżenia prawdopodobieństwa  $p_h$  hipotezy o dominacji reszki. Rysunek pokazuje także pozycję wyniku obliczonego na podstawie częstościowej interpretacji prawdopodobieństwa, to jest wartości  $p_h = 399/700 = 0.570$ . Nie wiadomo, dlaczego akurat ta wartość miałaby reprezentować dokładną, ale nieznaną, ze względu na niedostatek danych, wartość prawdopodobieństwa hipotezy. Wartość sugerowana przez częstościową interpretację nie jest ani bardziej ani mniej wiarygodna niż każda inna wartość prawdopodobieństwa z zakresu wartości możliwych  $p_h \in [399/921, 620/921]$ .

**Rozpatrzmy teraz sytuację trzecią** w której posiadamy pełny satysfakcjonujący zbiór dowodów o liczności  $n = n_{SEC} = 921$  składający się z  $n_h = 531$  reszek oraz z  $n_{\bar{h}} =$

390 orłów. Ponieważ posiadamy pełny satysfakcjonujący zbiór dowodowy to w oparciu o kompletnościową interpretację prawdopodobieństwa uzyskujemy wyniki podane poniżej.

$$p_h = n_h/n_{SEC} = 531/921 \text{ (reszka)}, \quad p_{\bar{h}} = n_{\bar{h}}/n_{SEC} = 390/921 \text{ (orzec)}$$

$$p_h + p_{\bar{h}} = (531/921) + (390/921) = 1$$

Identyczne wartości prawdopodobieństw uzyskamy w tym wypadku na podstawie częstościowej interpretacji prawdopodobieństwa bowiem wartość  $n$  całkowitej liczby prób, z jakiej korzysta ta interpretacja jest dokładnie równa satysfakcjonującej liczbie prób, tzn.  $n = n_{SEC} = 921$ . Wartość prawdopodobieństwa  $p_h = 531/921$  hipotezy o dominacji reszki jest bardzo bliska dokładnej wartości tego prawdopodobieństwa. Maksymalny błąd dowodu, zgodnie ze wzorem Chernoffa (4) nie przekracza 0.01. Idealnie dokładnej wartości prawdopodobieństwa nie można określić, gdyż wymagałoby to posiadania wyników z nieskończonej liczby prób rzutu monetą ( w ogólnym przypadku nieskończenie wielu dowodów). Uzyskane wyniki pokazane są na rys. 9.

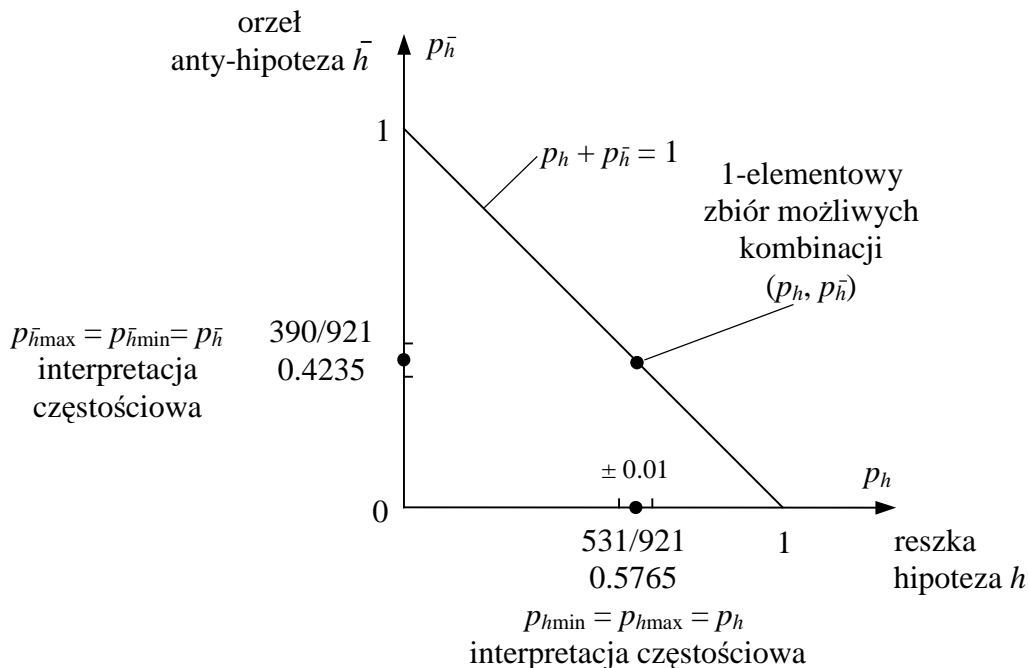
$$n = 921, \quad n_h = 531, \quad n_{\bar{h}} = 390, \quad n_h + n_{\bar{h}} = n_{SEC}$$

Wynik interpretacji kompletnościowej:

$$p_h \in \left[ \frac{n_h}{n_{SEC}}, 1 - \frac{n_{\bar{h}}}{n_{SEC}} \right] = \left[ \frac{531}{921}, 1 - \frac{390}{921} \right] = \left[ \frac{531}{921}, \frac{531}{921} \right] = [0.5765, 0.5765]$$

$$p_{\bar{h}} = 1 - p_h = 1 - \frac{531}{921} = \frac{390}{921} = 0.4235 \quad \text{możliwy błąd: } \epsilon \leq 0.01$$

$$\text{Wynik interpretacji częstościowej: } p_h = \frac{n_h}{n} = \frac{531}{921} \quad p_{\bar{h}} = \frac{390}{921}$$

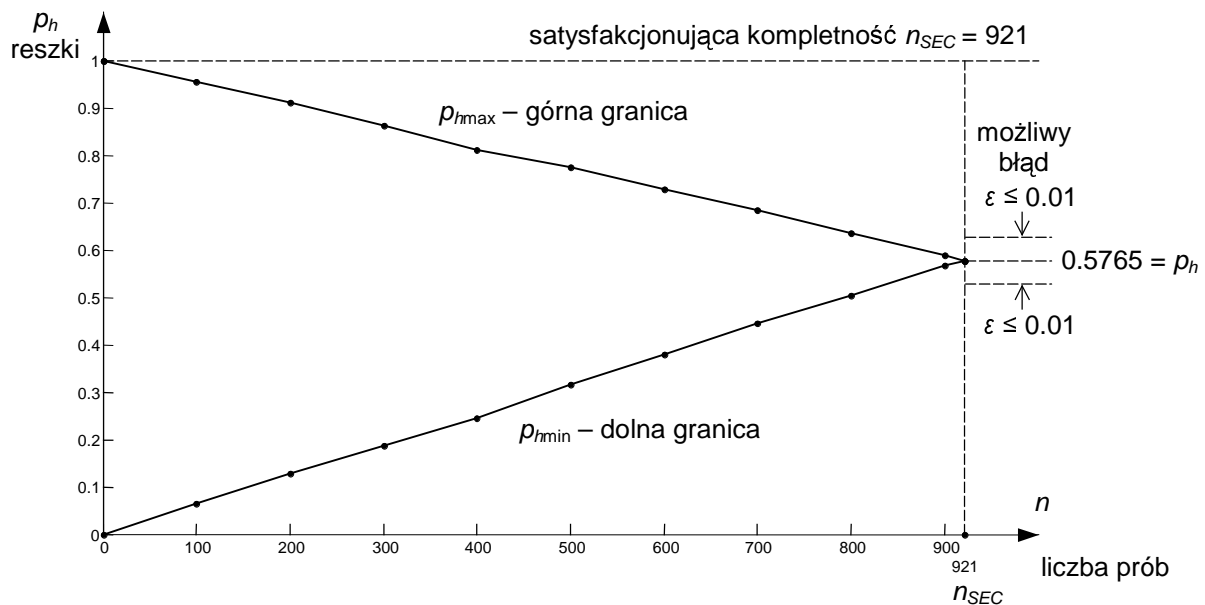


Rysunek 9: Ilustracja wyników określenia prawdopodobieństwa  $p_h$  hipotezy o przewadze reszki oraz prawdopodobieństwa  $p_{\bar{h}}$  anty-hipotezy o przewadze orła w sytuacji posiadania pełnego satysfakcjonującego zbioru ewidencyjnego  $n = n_{SEC} = 921$  (531 reszek i 390 orłów).

W tabeli 5 oraz na rys. 10 pokazana jest przykładowa historia przyrostu liczby dowodów w miarę kolejnych prób rzutu monetą, aż do uzyskania pełnej liczby dowodów  $n = n_{SEC} = 921$  wymaganej przez kompletność ewidencyjną SEC.

|               |   |        |        |        |        |        |        |        |        |        |                 |
|---------------|---|--------|--------|--------|--------|--------|--------|--------|--------|--------|-----------------|
| $n$           | 0 | 100    | 200    | 300    | 400    | 500    | 600    | 700    | 800    | 900    | $n_{SEC} = 921$ |
| $n_h$         | 0 | 59     | 117    | 172    | 226    | 292    | 349    | 409    | 464    | 522    | 531             |
| $n_{\bar{h}}$ | 0 | 41     | 83     | 128    | 174    | 208    | 251    | 291    | 336    | 378    | 390             |
| $p_{hmin}$    | 0 | 0.0641 | 0.1270 | 0.1868 | 0.2454 | 0.3170 | 0.3789 | 0.4441 | 0.5038 | 0.5668 | 0.5765          |
| $p_{hmax}$    | 1 | 0.9555 | 0.9099 | 0.8610 | 0.8111 | 0.7742 | 0.7275 | 0.6840 | 0.6352 | 0.5896 | 0.5765          |

Tabela 5: Wyniki z eksperymentu rzucania monetą aż do zrealizowania wszystkich  $n_{SEC} = 921$  prób wymaganych przez satysfakcjonującą kompletność dowodowa SEC. Oznaczenia:  $n$  – całkowita liczba prób,  $n_h$  – liczba reszek,  $n_{\bar{h}}$  – liczba orłów,  $p_{hmin}$  – minimalne prawdopodobieństwo hipotezy o dominacji reszki,  $p_{hmax}$  – maksymalne prawdopodobieństwo hipotezy.



Rysunek 10: Zmiana prawdopodobieństwa  $p_{hmin}$  hipotezy o dominacji reszki oraz prawdopodobieństwa  $p_{hmin}$  anty-hipotezy o dominacji orła w miarę przyrostu wyników prób rzutu monetą.

Rys. 10 pokazuje, jak zmieniają się obydwa prawdopodobieństwa, minimalne  $p_{hmin}$  oraz maksymalne  $p_{hmax}$  w miarę przyrostu sumarycznej liczby dowodów  $n$  i jak zmniejsza się luka między nimi ( $p_{hmax}p_{hmax}$ ) stanowiąca zakres niepewności identyfikowanego prawdopodobieństwa  $p_h$ .

Wartość prawdopodobieństwa  $p_h = 531/921 = 0.5765$  pokazana na rys. 10 nie jest całkowicie dokładną wartością prawdopodobieństwa hipotezy  $h$  (dominacja reszki) bowiem jej określenie wymagałoby przeprowadzenia nieskończonej liczby prób. Jest to jednak satysfakcjonujące przybliżenie, którego błąd, zgodnie ze wzorem Chernoffa (4) nie przekracza wartości 0.01.

## 6 Optymalna reprezentacja $p_{hR}$ zakresu niepewności $[p_{hmin}, p_{hmax}]$ prawdopodobieństwa $p_h$ hipotezy

W zadaniach realnego podejmowania decyzji często potrzebna jest jedno-liczbowa reprezentacja (singleton-representation) przedziału niepewności, bowiem taka właśnie reprezentacja najbardziej ułatwia praktykom oraz niespecjalistom podejmowanie decyzji. Aby określić jedno-liczbową reprezentację dla przedziału niepewności należy odpowiedzieć na pytanie: “Która wartość  $p_{hR}$  prawdopodobieństwa leżąca w rozpatrywanym przedziale niepewności  $[p_{hmin}, p_{hmax}]$  najlepiej reprezentuje ten przedział?” Aby odpowiedzieć na to pytanie należy sformułować kryterium oceny. Dobrym kryterium jest tu kryterium (18), gdzie  $p_{hR}^*$  oznacza wartość ‘kandydującą’ na reprezentację  $p_{hR}$ .

$$p_{hR} = \min[\max(p_{hR}^* - p_{hmin}, p_{hmax} - p_{hR}^*)] \quad (18)$$

$$p_{hR}^* \in [p_{hmin}, p_{hmax}]$$

Kryterium to minimalizuje maksymalny możliwy błąd reprezentacji  $p_{hR}$  względem dokładnej, nieznannej wartości prawdopodobieństwa  $p_h$  hipotezy  $h$ . Oznaczmy przez  $p_{hR}$  optymalną jedno-liczbową reprezentację przedziału niepewności  $[p_{hmin}, p_{hmax}]$  wartości  $p_h$  spośród wszystkich możliwych reprezentacji  $p_{hR}^*$  zawartych w tym przedziale. Łatwo sprawdzić, że optymalną reprezentacją jest średnia wartość (19) ograniczeń  $p_{hmin}$  oraz  $p_{hmax}$  przedziału.

$$p_{hR} = 0.5(p_{hmin} + p_{hmax}) \quad (19)$$

Konieczna jest przy tym **bardzo ważna uwaga: optymalna reprezentacja  $p_{hR}$  zwykle nie jest dokładną wartością prawdopodobieństwa  $p_h$**  (choćby niekiedy może nią być) bowiem prawdopodobieństwo to nie może być dokładnie poznane (wymagałoby to nieskończenie wielkiej liczby prób). Optymalna reprezentacja jest jedynie najlepszym szacunkiem tej wartości określonym na podstawie takiej liczby prób jaką aktualnie posiadamy. Jest szacunkiem który możemy przyjąć w celu podjęcia decyzji związanej z rozpatrywanym problemem. Stosowanie tej reprezentacji w warunkach częściowej niewiedzy zapobiega popełnianiu dużych i bardzo dużych błędów w rozwiązywanych problemach. Optymalna reprezentacja  $p_{hR}$  może być przedstawiona w bardziej szczegółowej formie (20) po wstawieniu do niej wzorów (14) i (15).

$$p_{hR} = 0.5(p_{hmin} + p_{hmax}) = 0.5 + 0.5(n_h - n_{\bar{h}})/n_{SEC} \quad (20)$$

Analiza powyższego wzoru nasuwa ciekawe spostrzeżenia przedstawione poniżej.

Kompletnościowa estymata  $p_{hR}$  prawdopodobieństwa zależy liniowo zarówno od liczby dowodów  $n_h$  potwierdzających hipotezę  $h$  jaki i liczby dowodów  $n_{\bar{h}}$  potwierdzających anty-hipotezę  $\bar{h}$ . Każde pojedyncze potwierdzenie czy to hipotezy czy anty-hipotezy zmienia kompletnościową estymatę  $p_{hR}$  o taką samą wartość bezwzględną równą  $0.5/n_{SEC}$  (potwierdzenie hipotezy zwiększa estymatę  $p_{hR}$ , zaś potwierdzenie anty-hipotezy zmniejsza ją). **Oznacza to, że wartość dowodowa każdego pojedynczego dowodu (wyniku rzutu monetą) jest identyczna.** Natomiast w przypadku estymaty częstościowej tak nie jest. Wynika to wyraźnie ze wzoru:

$$fr_h = n_h/n = n_h/(n_h + n_{\bar{h}}), \quad (21)$$

proponowanego przez częstościową interpretację prawdopodobieństwa, gdzie oznaczenie  $fr_h$  oznacza częstość (frequency) względną potwierżeń hipotezy  $h$  w sumarycznej liczbie



dowodów  $n$ . Częstość  $fr_h$  będzie w dalszym ciągu stosowana jako oznaczenie dla odróżnienia częstościowej estymaty prawdopodobieństwa od estymaty kompletnościowej  $p_{hr}$  oraz od dokładnej wartości prawdopodobieństwa hipotezy  $p_h$ . Zależność względnej częstości  $fr_h$  zarówno od liczby potwierdzeń  $n_h$  hipotezy  $h$  jak i liczby potwierdzeń  $n_{\bar{h}}$  anty-hipotezy  $\bar{h}$  jest nieliniowa (zarówno  $n_h$  jak i  $n_{\bar{h}}$  występuje w mianowniku wzoru (21)). Oznacza to, że pojedyncze potwierdzenie hipotezy  $h$  zmienia częstościową estymatę  $fr_h$  prawdopodobieństwa  $p_h$  o inną wartość bezwzględną aniżeli pojedyncze potwierdzenie anty-hipotezy  $\bar{h}$ . Podobnie, w przypadku dwóch kolejnych potwierdzeń hipotezy  $h$  potwierdzenie pierwsze zmienia estymatę częstościową o inną wartość liczbową aniżeli potwierdzenie drugie. Oznacza to, że znaczenie pojedynczych dowodów (pojedynczych rzutów moneta) jest niejednakowe, co jest nielogiczne i nieuzasadnione. Uwagę na ten fakt zwracali wielokrotnie różni znani naukowcy, np. Burdzy [1] i Hajek [7].

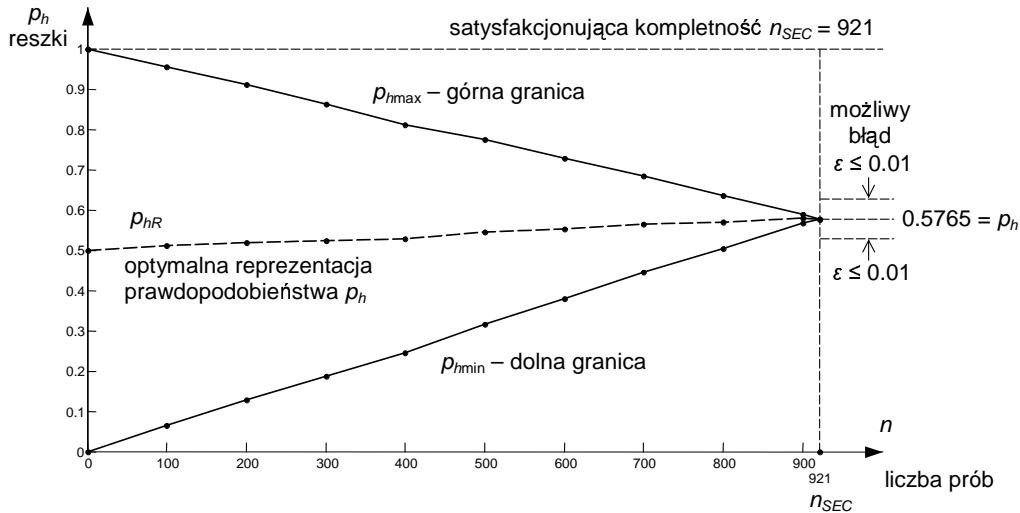
Interesującą kwestią jest jak optymalna reprezentacja  $p_{hR}$  zmienia się wraz ze zwiększającą się liczbą  $n$  dowodów (wyników prób) dla  $n \leq n_{SEC}$ . W tabeli 6 podane są przykładowe wyniki eksperymentów rzucania monetą.

| $n$             | 0   | 100    | 200    | 300    | 400    | 500    | 600    | 700    | 800    | 900    | $n_{SEC} = 921$ |
|-----------------|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|-----------------|
| $n_h$           | 0   | 59     | 117    | 172    | 226    | 292    | 349    | 409    | 464    | 522    | 531             |
| $n_{\bar{h}}$   | 0   | 41     | 83     | 128    | 174    | 208    | 251    | 291    | 336    | 378    | 390             |
| $p_{hmin}$      | 0   | 0.0641 | 0.1270 | 0.1868 | 0.2454 | 0.3170 | 0.3789 | 0.4441 | 0.5038 | 0.5668 | 0.5765          |
| $p_{hr}$        | 0.5 | 0.5098 | 0.5185 | 0.5239 | 0.5282 | 0.5456 | 0.5532 | 0.5641 | 0.5695 | 0.5782 | 0.5765          |
| $p_{hmax}$      | 1   | 0.9555 | 0.9099 | 0.8610 | 0.8111 | 0.7742 | 0.7275 | 0.6840 | 0.6352 | 0.5896 | 0.5765          |
| $\frac{n_h}{n}$ | –   | 0.5900 | 0.5850 | 0.5733 | 0.5650 | 0.5840 | 0.5817 | 0.5843 | 0.5800 | 0.5800 | 0.5765          |
| $ \Delta $      | –   | 0.0802 | 0.0665 | 0.0494 | 0.0368 | 0.0384 | 0.0285 | 0.0202 | 0.0105 | 0.0018 | 0               |

Tabela 6: Przykładowe wyniki eksperymentów rzucania monetą. Oznaczenia:  $n$  – sumaryczna liczba rzutów,  $n_h$  – liczba reszek,  $n_{\bar{h}}$  – liczba orłów,  $p_{hmin}$  – dolna granica prawdopodobieństwa  $p_h$  hipotezy,  $p_{hmax}$  – górna granica prawdopodobieństwa hipotezy,  $p_{hR}$  – optymalna reprezentacja zakresu niepewności prawdopodobieństwa  $p_h$  (estymata tego prawdopodobieństwa),  $n_h/n = fr_h$  – wartość prawdopodobieństwa hipotezy obliczona według częstościowej interpretacji prawdopodobieństwa,  $\Delta = p_{hR} - n_h/n$  – różnica między estymatą prawdopodobieństwa określoną na podstawie interpretacji kompletnościowej i częstościowej.

Na rys. 11 pokazano zmiany dolnej i górnej granicy prawdopodobieństwa  $p_h$  oraz optymalnej reprezentacji  $p_{hR}$  tego zakresu czyli optymalnej estymaty kompletnościowej.

Jak pokazuje rys. 11, przy małej liczbie  $n$  dowodów (wyników prób) niepewność ( $p_{hmax} - p_{hmin}$ ) prawdopodobieństwa  $p_h$  jest bardzo duża, jednak ze wzrostem liczby dowodów stopniowo maleje do minimum (do wartości możliwego błędu  $\epsilon = 0.01$  w sensie wzoru Chernoffa (4)). Także kompletnościowa estymata  $p_{hR}$  prawdopodobieństwa  $p_h$  w miarę przybywania dowodów stopniowo i bez fluktuacji dąży do swej końcowej wartości  $0.5765 = 531/921$ . Wiarygodność obliczonej wartości  $p_h = 531/921$  wynosi 0.99 co oznacza, że gdybyśmy powtarzali serię 921 rzutów 100 razy, to tylko w jednej serii na 100 wartość  $p_h$  obliczona na podstawie wyników serii będzie różnić się od dokładnej, prawdziwej wartości tego prawdopodobieństwa o więcej niż 0.01. W tabeli 6 oraz na rys. 11 przedstawiono przebieg zmian estymaty kompletnościowej  $p_{hR}$  dla dużych liczb dowodów aż do  $n = 921$ . Natomiast w tabeli 7 przedstawiono kolejne wyniki przykładowej serii 10



Rysunek 11: Ilustracja procesu zmniejszania się zakresu niepewności  $[p_{hmin}, p_{hmax}]$  prawdopodobieństwa  $p_h$  hipotezy oraz towarzyszących temu zmian kompletnościowej estymaty  $p_{hR}$  tego prawdopodobieństwa w miarę powiększania się liczby posiadanych dowodów  $n$  (wyników rzutów monetą).

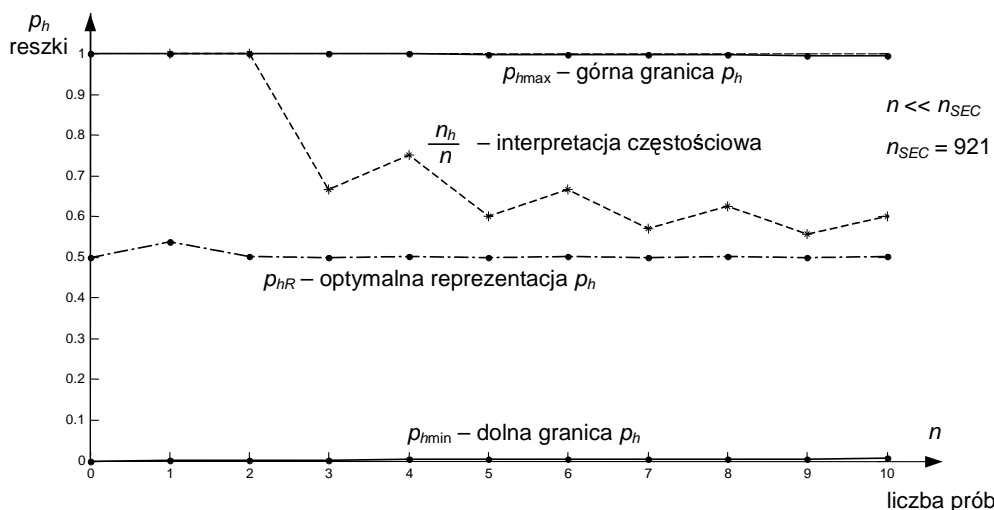
rzutów monetą wraz z wartościami obliczonych estymat prawdopodobieństwa.

| $n$             | 0   | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      | 10     |
|-----------------|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $n_h$           | 0   | 1      | 2      | 2      | 3      | 3      | 4      | 4      | 5      | 5      | 6      |
| $n_{\bar{h}}$   | 0   | 0      | 0      | 1      | 1      | 2      | 2      | 3      | 3      | 4      | 4      |
| $p_{hmin}$      | 0   | 0.0011 | 0.0022 | 0.0022 | 0.0033 | 0.0033 | 0.0043 | 0.0043 | 0.0054 | 0.0054 | 0.0065 |
| $p_{hr}$        | 0.5 | 0.5385 | 0.5011 | 0.5005 | 0.5011 | 0.5005 | 0.5011 | 0.5005 | 0.5011 | 0.5005 | 0.5011 |
| $p_{hmax}$      | 1   | 1      | 1      | 0.9989 | 0.9989 | 0.9978 | 0.9978 | 0.9967 | 0.9967 | 0.9956 | 0.9956 |
| $\frac{n_h}{n}$ | –   | 1      | 1      | 0.6667 | 0.7500 | 0.6000 | 0.6667 | 0.5714 | 0.6250 | 0.5555 | 0.6000 |
| $ \Delta $      | –   | 0.4615 | 0.4999 | 0.1662 | 0.2498 | 0.0995 | 0.1656 | 0.0709 | 0.1239 | 0.0550 | 0.0989 |

Tabela 7: Wyniki kolejnych 10 rzutów monetą i odpowiadające im wartości dolnej  $p_{hmin}$  i górnej  $p_{hmax}$  granicy estymowanego prawdopodobieństwa  $p_h$ , wartości estymaty kompletnościowej  $p_{hR}$ , estymaty częstościowej  $fr_h = n_h/n$  oraz bezwzględnej różnicy obydwu estymat  $|\Delta| = |p_{hR} - n_h/n$ .

Na rys. 12 przedstawiono wizualnie przebieg zmian porównywanych estymat  $p_{hR}$  oraz  $fr_h$  dla małej liczby dowodów  $n \leq 10$  podanych w tabeli 7.

Rys. 12 uzmysławia jak mało wiarygodna jest estymata częstościowa przy małej liczbie  $n$  dowodów. Jeżeli posiadamy tylko 1 dowód ( $n = 1$ , ‘single case problem’) to estymata częstościowa może mieć albo wartość 1, jak w przykładzie na rys. 12 albo wartość 0. Oznacza to, że estymata ta na podstawie jednego tylko dowodu sugeruje prawdopodobieństwo mające charakter pewności, co jest wnioskiem nadzwyczaj pochopnym. Przy wzroście liczby  $n$  dowodów estymata częstościowa dąży wprawdzie stopniowo do dokładnej wartości prawdopodobieństwa  $p_h$  hipotezy, jednak jej przebieg wykazuje znaczne fluktuacje, co oznacza, że kolejno dochodzące dowody (wyniki rzutów monetą) dość znacznie zmieniają szacowane wartości prawdopodobieństwa. Interpretacja częstościowa nie jest też w stanie



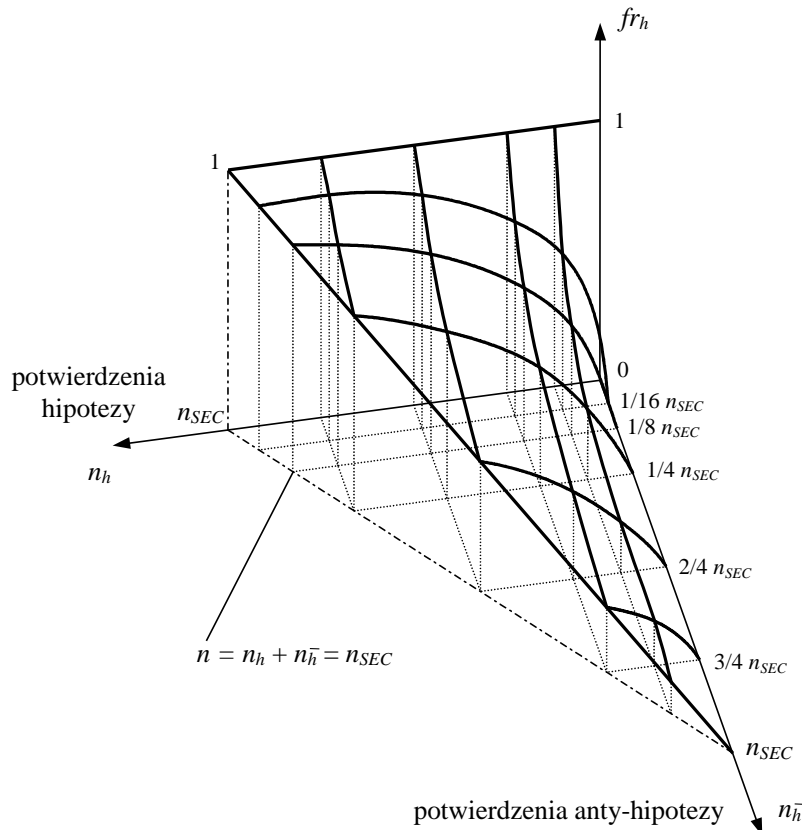
Rysunek 12: Przebieg zmian kompletnościowej estymaty  $p_{hR}$  oraz estymaty częstościowej  $fr_h = n_h/n$  w zależności od liczby dowodów dla przypadku małej liczby dowodów  $n \leq 10$ .

określić żadnej wartości prawdopodobieństwa dla zerowej liczby dowodów (w sytuacji braku jakichkolwiek dowodów). Mamy bowiem wówczas do czynienia z dzieleniem przez zero – wynik takiej operacji jest nieokreślony. Natomiast estymata kompletnościowa podaje wiarygodny i rozsądny wynik zarówno dla zerowej liczby dowodów ( $p_{hR} = 0.5$ ), dla przypadku jednego dowodu oraz dla każdej następnej liczby dowodów  $n$ . Przebieg tej estymaty wykazuje minimalne fluktuacje, daleko mniejsze niż w przypadku estymaty częstościowej. Estymata kompletnościowa przypomina więc osobę, która nie zmienia silnie swych opinii z każdym pojedynczym, następnym dowodem, lecz czyni to z rozmysłem. Estymata ta także stopniowo, krok po kroku zbliża się do dokładnej wartości prawdopodobieństwa  $p_h$ , co pokazane jest na rys. 11. Przy dużej liczbie  $n$  dowodów obydwie estymaty są do siebie bardzo zbliżone, z tym zastrzeżeniem, że estymata częstościowa może wykazywać znaczne fluktuacje również przy dużych wartościach  $n$ , natomiast estymata kompletnościowa fluktuacji takich nie wykazuje.

Na rys. 13 przedstawiono w widoku z przodu powierzchnię funkcyjną zależności prawdopodobieństwa hipotezy rozumianego w sensie częstościowej interpretacji prawdopodobieństwa jako względna częstość  $fr_h = n_h/(n_h + n_{\bar{h}})$ .

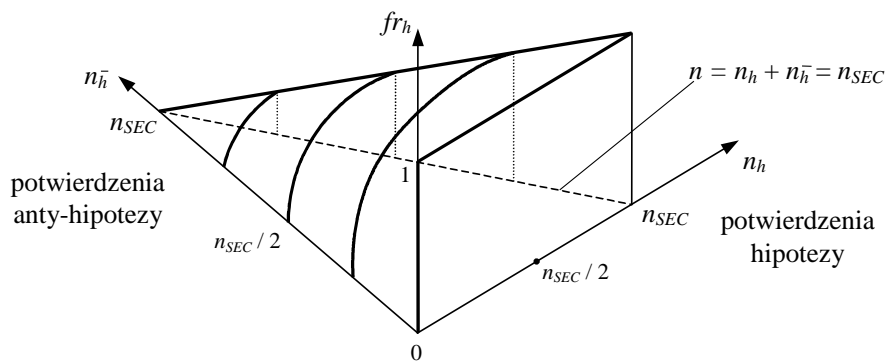
Zakrzywiona, nieliniowa powierzchnia zależności  $fr_h = f(n_h, n_{\bar{h}})$  dobrze wyjaśnia dlaczego kolejne potwierdzenia hipotezy  $h$  (kolejne reszki) nie mają jednakowej istotności dowodowej. Zależnie od tego jaka jest początkowa sytuacja dowodowa (aktualna liczba dowodów  $n_h$  hipotezy oraz dowodów  $n_{\bar{h}}$  anty-hipotezy dodanie kolejnego pojedynczego dowodu powoduje różne zmiany częstości  $fr_h$ , co świadczy o niejednakowym traktowaniu identycznych dowodów przez częstościową reprezentację prawdopodobieństwa. Jak zobaczymy dalej, kompletnościowa reprezentacja traktuje wszystkie dowody w jednakowy sposób przykładając do nich identyczną wagę, niezależnie od tego jaki jest początkowy stan dowodowy czyli początkowe wartości  $n_h$  oraz  $n_{\bar{h}}$ .

Rysunek 14 dobrze obrazuje przyczynę nielogiczności wyników obliczania prawdopodobieństwa w sensie względnej częstości  $fr_h$ . Powierzchnia funkcyjna zależności  $fr_h = f(n_h, n_{\bar{h}})$  jest powierzchnią nieliniową posiadającą strefy nieczułości na swoich brzegach. I tak, jeśli brak jest dowodów na poparcie anty-hipotezy  $\bar{h}$  ( $n_{\bar{h}} = 0$ ) a istnieją dowody na poparcie hipotezy  $h$  ( $n_h > 0$ ) to niezależnie od tego jak duża jest liczba tych dowo-



$$p_h = \frac{n_h}{n} = \frac{n_h}{n_h + n_{\bar{h}}} \quad p_{\bar{h}} = \frac{n_{\bar{h}}}{n} = \frac{n_{\bar{h}}}{n_h + n_{\bar{h}}}$$

Rysunek 13: Powierzchnia funkcyjna zależności częstości względnej  $fr_h = n_h/n = n_h/(n_h + n_{\bar{h}})$  od liczby potwierdzeń  $n_h$  hipotezy  $h$  oraz liczby potwierdzeń  $n_{\bar{h}}$  anty-hipotezy  $\bar{h}$ . Widok z przodu.



$$fr_h = \frac{n_h}{n} = \frac{n_h}{n_h + n_{\bar{h}}} \quad fr_{\bar{h}} = \frac{n_{\bar{h}}}{n} = \frac{n_{\bar{h}}}{n_h + n_{\bar{h}}}$$

Rysunek 14: Powierzchnia funkcyjna zależności częstości względnej  $fr_h = n_h/n = n_h/(n_h + n_{\bar{h}})$  hipotezy  $h$  od liczby  $n_h$  potwierdzeń hipotezy oraz od liczby  $n_{\bar{h}}$  potwierdzeń anty-hipotezy  $\bar{h}$  w widoku z tyłu,  $n_{SEC}$  – całkowita liczba prób wymagana przez kompletność ewidencyjną SEC.

dów, czy  $n_h$  wynosi 1, czy 2, czy 10, czy 100, prawdopodobieństwo częstościowe hipotezy  $f_{r_h} = 1$  i nie zmienia się z liczbą dowodów  $n_h$ . Nie widać tu żadnego związku między liczbą potwierdzeń hipotezy a jej prawdopodobieństwem. A przecież logicznym jest, że ze zmianą liczby potwierdzeń hipotezy jej prawdopodobieństwo powinno się zmieniać, że nie powinno ono być stałe i niezmiennie. Tak więc częstość względna nie odwzorowuje prawdopodobieństwa  $p_h$  hipotezy przy braku potwierdzeń ( $n_{\bar{h}} = 0$ ) anty-hipotezy. Jak zobaczymy dalej, kompletnościowa interpretacja prawdopodobieństwa wady tej nie posiada i w sposób logiczny informuje o prawdopodobieństwie hipotezy także w tym przypadku.

Na rys. 15 przedstawione są jednocześnie 3 zależności funkcyjne wynikające z kompletnościowej interpretacji prawdopodobieństwa:  $p_{hmin} = f_1(n_h, n_{\bar{h}})$ ,  $p_{hmax} = f_3(n_h, n_{\bar{h}})$ , oraz estymata kompletnościowa  $p_{hR} = f_2(n_h, n_{\bar{h}})$ . Na rysunku tym szczególnie dobrze widać zbieganie się powierzchni dolnego i górnego ograniczenia prawdopodobieństwa hipotezy  $h$  oraz jej optymalnej estymaty. Zbieganie tych powierzchni w jedną linię następuje przy liczbie dowodów  $n = n_h + n_{\bar{h}} = n_{SEC}$  wymaganej przez satysfakcjonujący zbiór dowodowy SEC. Wówczas dolna i górna granica prawdopodobieństwa hipotezy  $h$  sumuje się do jedności niezależnie od tego jaki jest stosunek ilościowy liczby potwierdzeń hipotezy  $n_h$  oraz anty-hipotezy  $n_{\bar{h}}$ .

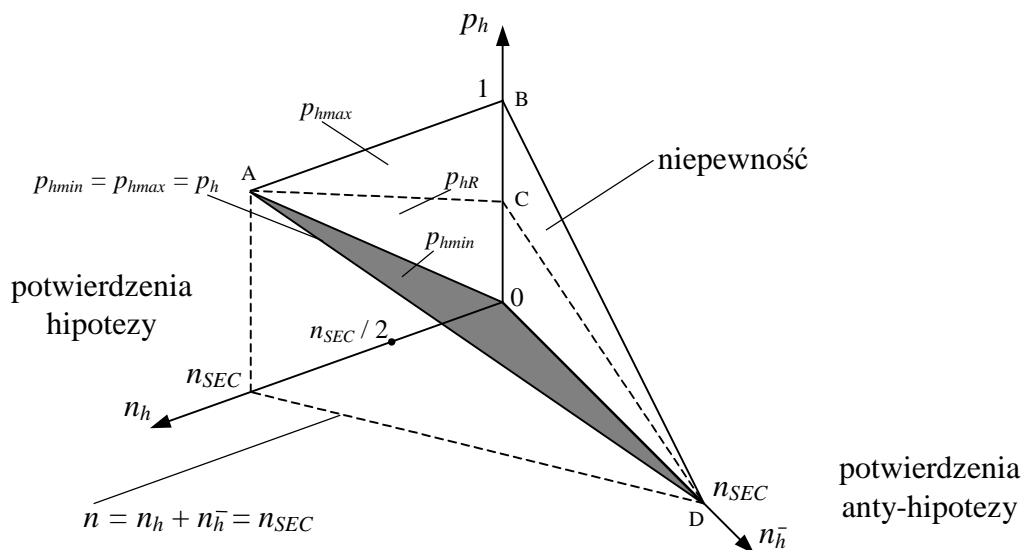
$$p_{hmin} + p_{hmax} = 1$$

Jeżeli sumaryczna liczba wszystkich dowodów  $n_h + n_{\bar{h}} = n < n_{SEC}$  to wartości dolnego i górnego ograniczenia prawdopodobieństwa nie sumują się do jedności a istniejąca między nimi różnica ( $p_{hmax} - p_{hmin}$ ) oznacza niepewność naszej wiedzy o prawdziwej wartości prawdopodobieństwa hipotezy  $p_h$ . Na rys. 16 przedstawiono te same powierzchnie funkcyjne w widoku z tyłu.

Rysunek ten dobrze obrazuje ocenę prawdopodobieństwa hipotezy przez interpretację kompletnościową przy braku dowodów  $n_{\bar{h}} = 0$  potwierdzających anty-hipotezę i przy istnieniu (bądź nieistnieniu) dowodów potwierdzających hipotezę ( $n_h \geq 0$ ). Ze wzrostem liczby potwierdzeń  $n_h$  hipotezy rośnie minimalna wartość  $p_{hmin}$  hipotezy (co jest logiczne), nie zmienia się natomiast górna wartość ograniczenia  $p_{hmax}$  co wynika z kolei logicznie z faktu, że liczba potwierdzeń anty-hipotezy  $n_{\bar{h}} = 0$  nie zmienia się (jest stała i równa 0).

Brzegowe, specjalne sytuacje oceny prawdopodobieństwa jeszcze lepiej obrazują i wyjaśniają rys. 17 i rys. 18. Na rys. 17 przedstawiona jest porównanie ocen prawdopodobieństwa hipotezy  $p_h$  przez interpretację kompletnościową i częstościową w sytuacji specjalnej, gdy istnieją dowody potwierdzające hipotezę ( $n_h \geq 0$ ) ale brak jest jakichkolwiek dowodów potwierdzających anty-hipotezę, to znaczy  $n_{\bar{h}} = 0$ . Przykładem takiej sytuacji jest sytuacja lekarza, który bada prawdziwość hipotezy  $h$ : 'regularne uprawianie sportu zapobiega otyłości' i który w swej bazie danych posiada dane np. 10 osób, które uprawiają sport i są szczupłe, potwierdzają więc hipotezę  $h$  lekarza ( $n_h = 10$ ). Jednak lekarz nie posiada na razie w swej bazie danych przykładów osób potwierdzających anty-hipotezę  $NOT h = \bar{h}$ , które mimo regularnego uprawiania sportu nie są szczupłe, a więc liczba potwierdzeń anty-hipotezy  $n_{\bar{h}} = 0$ . Czy lekarz na tej podstawie ma wnioskować że regularne uprawianie sportu na pewno i w każdym przypadku zapewnia szczupłość? To znaczy, czy ma wnioskować, że  $p_h = 1$  a  $p_{\bar{h}} = 0$ ? Powyższy przykład dobrze jest mieć przed oczami czytając podane w dalszym ciągu analizy.

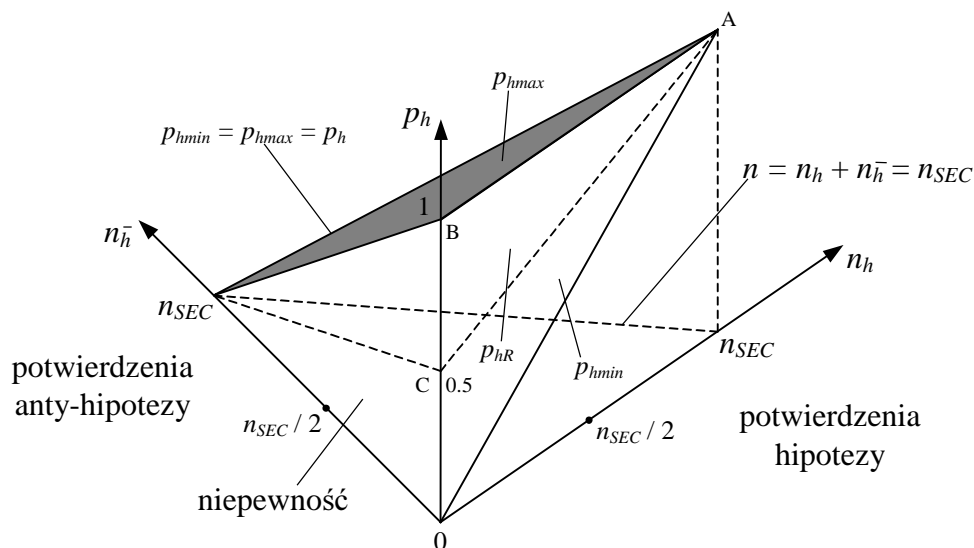
Jak pokazuje rys. 17a, kompletnościowa interpretacja wraz ze wzrostem liczby potwierdzeń  $n_h$  hipotezy w logiczny sposób generuje coraz wyższe wartości dolnego ograniczenia prawdopodobieństwa hipotezy  $p_{hmin}$  nie zmieniając jednocześnie wartości ograniczenia górnego  $p_{hmax}$ , które jest zależne od liczby potwierdzeń  $n_{\bar{h}}$  anty-hipotezy. Ponieważ liczba



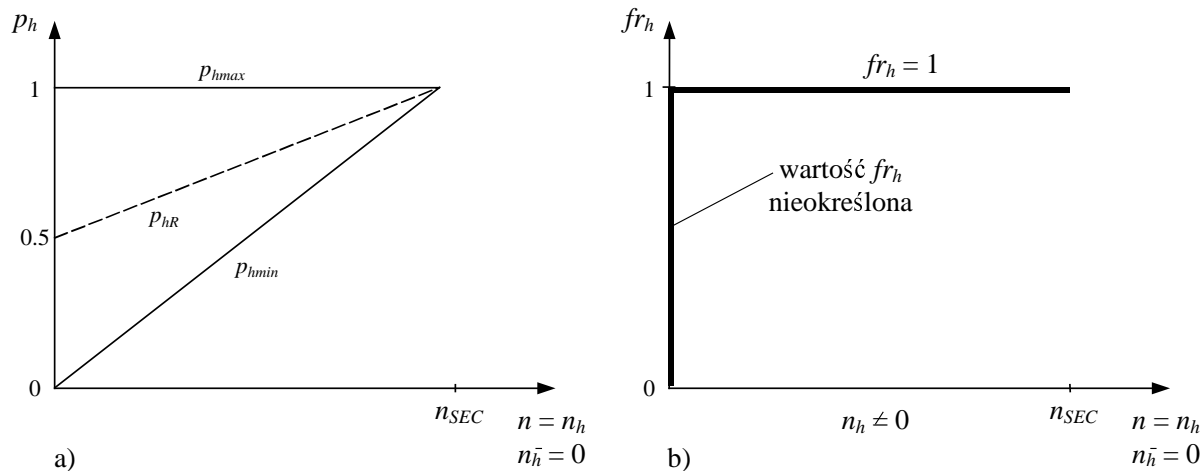
$$p_{hmin} = \frac{n_h}{n_{SEC}} \quad p_{hmax} = 1 - \frac{n_{\bar{h}}}{n_{SEC}}$$

$$p_{hR} = \frac{p_{hmin} + p_{hmax}}{2} = \frac{1}{2} + \frac{1}{2} \frac{n_h - n_{\bar{h}}}{n_{SEC}}$$

Rysunek 15: Powierzchnie zależności funkcyjnych prawdopodobieństwa hipotezy  $h$  generowanych przez kompletnościową interpretację prawdopodobieństwa w widoku z przodu, powierzchnia dolnego ograniczenia  $p_{hmin} = f_1(n_h, n_{\bar{h}})$ , powierzchnia optymalnej estymaty  $p_{hR} = f_2(n_h, n_{\bar{h}})$ , powierzchnia górnego ograniczenia  $p_{hmax} = f_3(n_h, n_{\bar{h}})$ .



Rysunek 16: Powierzchnie funkcyjne zależności wynikające z kompletnościowej interpretacji prawdopodobieństwa:  $p_{hmin} = f_1(n_h, n_{\bar{h}})$ ,  $p_{hmax} = f_3(n_h, n_{\bar{h}})$  oraz estymata kompletnościowa  $p_{hR} = f_2(n_h, n_{\bar{h}})$  w widoku z tyłu.



$$p_{hmin} = n_h/n_{SEC} \quad p_{hmax} = 1(n_h/n_{SEC})$$

$$p_{hR} = 0.5(p_{hmin} + p_{hmax}) = 0.5 + 0.5(n_h - n_{\bar{h}})/n_{SEC}$$

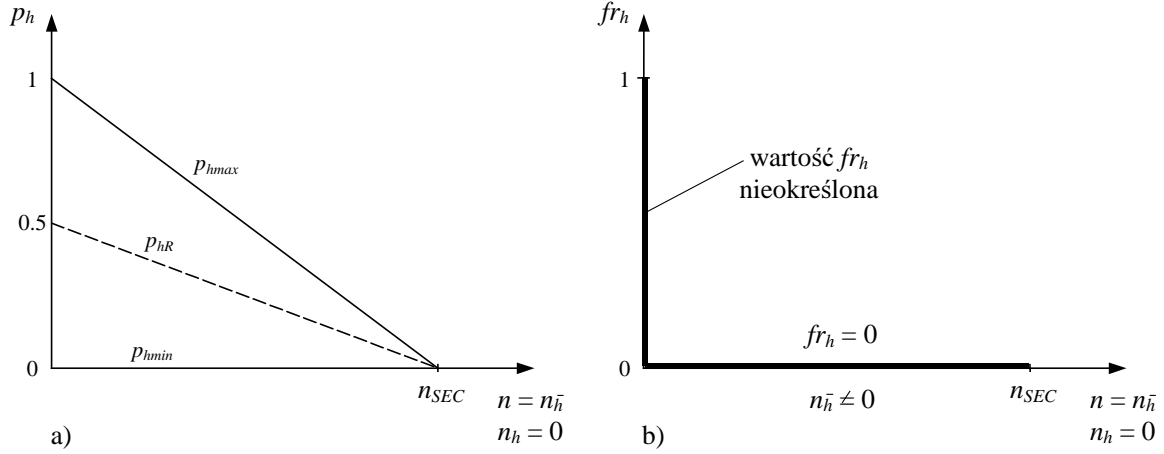
$$fr_h = n_h/n = n_h/(n_h + n_{\bar{h}})$$

Rysunek 17: Porównanie szacunków prawdopodobieństwa hipotezy  $p_h$  generowanych przez kompletnościową (a) i przez częstościową (b) interpretację prawdopodobieństwa w sytuacji istnienia potwierżeń hipotezy  $h$  ( $n_h \geq 0$ ) oraz całkowitego braku potwierżeń anty-hipotezy  $\bar{h}$  ( $n_{\bar{h}} = 0$ ).

ta jest stała i równa 0 wartość  $p_{hmax}$  też jest stała i równa 1, zgodnie ze wzorem podanym na rysunku. Natomiast częstościowa interpretacja pokazana na rys. 17b nie reaguje w ogóle na wzrost liczby potwierżeń  $n_h$  hipotezy  $h$  generując stałą wartość estymaty prawdopodobieństwa  $fr_h = 1$ , co jest nielogiczne. Dodatkowo, estymata częstościowa nie jest w stanie przypisać żadnej wartości prawdopodobieństwa dla stanu dowodowego  $n_h = 0$  oraz  $n_{\bar{h}} = 0$  bowiem mamy tu do czynienia ze stanem nieokreślonym dzielenia przez zero. Natomiast interpretacja kompletnościowa generuje dla tej sytuacji wiarygodną estymatę prawdopodobieństwa 0.5 (w sensie optymalnej reprezentacji przedziału niepewności, który jest tu maksymalny i równy 1).

Na rys. 18 przedstawiono porównanie prawdopodobieństw generowanych przez interpretację kompletnościową oraz częstościową dla sytuacji odwrotnej niż przedstawiona na rys. 17, w której istnieją jedynie dowody na potwierdzenie anty-hipotezy (dominacja orła), tzn.  $n_{\bar{h}} \geq 0$ , a brak jest całkowicie dowodów na poparcie hipotezy (dominacja reszki) tzn.  $n_h = 0$ . Innym przykładem o którym warto pamiętać czytając podane dalej analizy teoretyczne może być przykład policjanta badającego prawdziwość hipotezy  $h$ : „Kierowca jeżdzący bardzo szybko powoduje wypadki”. Załóżmy, że policjant na razie ma w swojej bazie danych przykłady 10 kierowców, którzy jeżdżą bardzo szybko i mimo to nie mieli dotychczas wypadku (są to przykłady popierające anty-hipotezę, a więc mamy  $n_{\bar{h}} = 10$ ) a nie posiada żadnych przykładów kierowców, którzy jeżdżą szybko i wywołali wypadki. Nie ma więc na razie przykładów potwierdzających hipotezę, tzn.  $n_h = 0$ . Czy policjant ma na tej podstawie wnioskować (jak sugeruje częstościowa interpretacja prawdopodobieństwa), że szybka jazda nigdy (z prawdopodobieństwem 1) nie powoduje wypadków?

Rysunek 18a pokazuje, że ponieważ liczba potwierżeń hipotezy  $n_h = 0$  (jest niezmienna i równa 0) to minimalne prawdopodobieństwo  $p_{hmin}$  tej hipotezy też jest również stałe i równe zero ( $p_{hmin} = 0$ ). Na początku, przy braku i przy małej liczbie potwierżeń  $n_{\bar{h}}$  anty-



$$p_{hmin} = n_h/n_{SEC} \quad p_{hmax} = 1(n_h/n_{SEC})$$

$$p_{hR} = 0.5(p_{hmin} + p_{hmax}) = 0.5 + 0.5(n_h - n_{\bar{h}})/n_{SEC}$$

$$fr_h = n_h/n = n_h/(n_h + n_{\bar{h}})$$

Rysunek 18: Porównanie szacunków prawdopodobieństw generowanych przez kompletnościową (a) i przez częstościową (b) interpretację prawdopodobieństwa w specjalnym przypadku istnienia jedynie dowodów potwierdzających anty-hipotezę ( $n_{\bar{h}}h \geq 0$ ) oraz całkowitego braku dowodów potwierdzających hipotezę ( $n_h = 0$ ).

hipotezy maksymalne prawdopodobieństwo hipotezy  $p_{hmax}$  jest wysokie. Jednak w miarę jak wzrasta liczba  $n_{\bar{h}}$  potwierdzeń anty-hipotezy, możliwe maksymalne prawdopodobieństwo hipotezy maleje, co jest logiczne, bowiem potwierdzenia anty-hipotezy zmniejszają potencjalne szanse hipotezy. Natomiast w przypadku interpretacji częstościowej przedstawionej na rys. 18b niezależnie od liczby potwierdzeń  $n_{\bar{h}}$  anty-hipotezy, które przecież zmniejszają szanse hipotezy, interpretacja ta sugeruje stałą, zerową wartość estymaty  $fr_h$  prawdopodobieństwa hipotezy. Zjawisko to można zinterpretować następująco: ‘ponieważ nie mam dowodów na poparcie hipotezy to całkowicie wykluczam nawet częściową jej prawdziwość’.

## 7 Podsumowanie

W wykładzie przedstawiono nową kompletnościową interpretację prawdopodobieństwa, która nie posiada szeregu wad najczęściej obecnie stosowanej interpretacji częstościowej. W szczególności, interpretacja kompletnościowa w wiarygodny sposób modeluje prawdopodobieństwo dla pewnych krytycznych i specjalnych sytuacjach, takich jak mała liczba próbek (dowodów) oraz braku dowodów potwierdzających prawdziwość jednej z hipotez. Interpretacja częstościowa nie potrafi w tych sytuacjach podać akceptowalnych wartości prawdopodobieństwa. A właśnie z przypadkiem małej liczby próbek bądź też ich brakiem (luki informacyjne [20]) często spotykamy się w problemach rzeczywistych. Stąd zakres zastosowania interpretacji kompletnościowej jest znacznie większy niż częstościowej, którą można stosować dopiero przy większej liczbie próbek. Kompletnościowa interpretacja prawdopodobieństwa uzmysławia nam, że w wielu wypadkach nie można poznać jego dokładnej wartości a jedynie jego dolną i górną granicę. Oznacza to częściową niepewność i nieokreśloność prawdopodobieństwa. W wykładzie przedstawiono interpretację



kompletnościową w wersji dla najprostszego przypadku binominalnego, to znaczy przypadku dwóch hipotez  $H = \{h, NOT h\}$ . Wychodząc z tego przypadku interpretację kompletnościową można rozszerzyć na przypadek trinominalny, tetranominalny, quintonominalny, . . . , multinomialny.

## Literatura

- [1] Burdzy K. (2009) *The search for certainty. On the clash of science and philosophy of probability*. World Scientific, New Jersey.
- [2] Carnap. (1950) *Logical foundations of probability*. University Press, Chicago.
- [3] Chernoff H. (1952) *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*. Annals of Mathematical Statistics, 23 (4), pp. 493-507.
- [4] Dubois D., Prade H. (1988) *Possibility theory*. Plenum Press, New York and London.
- [5] de Finetti B. (1975) *Theory of probability. A critical introductory treatment*. Wiley, London.
- [6] <http://www.math.washington.edu/burdzy>
- [7] Stanford Encyclopedia of Philosophy. Free access by <http://plato.stanford.edu/entries/probability-interpret/>
- [8] [http://en.wikipedia.org/wiki/Category:Probability\\_interpretations](http://en.wikipedia.org/wiki/Category:Probability_interpretations)
- [9] [http://en.wikipedia.org/wiki/Chernoff\\_bound](http://en.wikipedia.org/wiki/Chernoff_bound)
- [10] <http://search4certainty.blogspot.com/>
- [11] Khrennikov A. (1999) *Interpretations of probability*. Brill Academic Pub., Utrecht, Boston.
- [12] Klir G.J, Bo Yuan (editors) (1996) *Fuzzy sets, fuzzy logic, and fuzzy systems*. Selected papers by Lofti A. Zadeh. World Scientific, Singapore, New Jersey, London, Hong Kong.
- [13] Laplace P.S. (1951) *A philosophical essay on probabilities*. English edition, New York, Dover Publication Inc.
- [14] Larose D.T. (2010) *Discovering statistics*. W.H. Freeman and Company, New York.
- [15] von Mises R. (1957) *Probability, statistics and the truth*. Dover, New York, Macmillan, second revised English edition.
- [16] Piegat A. (2010) *Uncertainty of probability*. Wykład plenarny na konferencji Instytutu Badań Systemowych Polskiej Akademii Nauk Ninth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, Warszawa, 8 październik 2010, <http://www.ibspan.waw.pl/ifs2010>, publikacja materiałów konferencji w roku 2011.
- [17] Popper K.R. (1957) *The propensity interpretation of the calculus of probability and the quantum theory*. In S. Koerner (editor), The Colston Papers, pp. 65-70.

- [18] Rocchi P. (2003) *The structural theory of probability: new ideas from computer science on the ancient problem of probability interpretation*. Kluwer Academic/Plenum Publishers, New York.
- [19] Shafer G. (1976) *A mathematical theory of evidence*. Princetown University Press, Princetown and London.
- [20] Yakov B.H. (2006) *Info-gap decision theory*. Second edition. Elsevier, Oxford, Amsterdam.
- [21] Zadeh L. A. (1965) *Fuzzy sets*. Information and Control, vol. 8, pp. 338-353.