Bayesowskie podejście do problemu niemieckich czołgów*

A Bayesian Treatment of the German Tank Problem

Cory M. Simon**

Problem niemieckich czołgów (PNC) (ang. *German tank problem*) ma ciekawe tło historyczne. Jest to także zajmujący problem statystycznej estymacji nadający się na zajęcia w szkole. Celem jest estymacja liczby czołgów numerowanych kolejnymi numerami serii na podstawie losowej próbki. W tym artykule wprowadzamy zarys bayesowskiego podejścia do PNC . Rozwiązanie przypisuje prawdopodobieństwo wystąpienia każdej możliwej liczby czołgów, co pozwala na ilościową ocenę niepewności estymacji (tj. średniej, wariancji itd. – przyp. tłum.). Daje to także możliwość wprowadzenia do rachunku wcześniejszej wiedzy lub przekonania o liczbie czołgów. Podamy tu odpowiednie przykłady. Na zakończenie dokonamy przeglądu podobnych zagadnień.

Tło historyczne

W celu sformułowania strategii wojskowej w czasie II wojny światowej (1939-1945), Alianci starali się ocenić tempo produkcji różnych typów sprzętu wojskowego wroga (czołgi, opony, rakiety itd.) Zwykłe metody oceny produkcji zbrojeniowej, włączając w to ekstrapolacje danych o przedwojennej produkcji, źródła wywiadowcze czy też przesłuchania jeńców wojennych, były w większości przypadków niewiarygodne lub sprzeczne. W 1943 roku amerykańskie i angielskie agencje wywiadu ekonomicznego wykorzystały niemiecką praktykę, stosowaną przy produkcji sprzętu, do statystycznej oceny produkcji uzbrojenia w Niemczech. Otóż niemieccy producenci oznaczali wyprodukowany sprzęt numerem seryjnym, a także kodem zawierającym datę i miejsce produkcji. Miało to ułatwiać zarządzanie częściami zamiennymi oraz znajdowanie producenta wadliwego sprzętu czy też części zamiennych, który mógłby dokonać kontroli jakości. Jednocześnie te liczby i kody na zdobytym sprzęcie nieprzyjaciela dawały Aliantom informację o produkcji uzbrojenia w III Rzeszy.

W celu oszacowania tempa produkcji czołgów, Alianci zbierali numery seryjne znajdujące się na podwoziach, silnikach, skrzyniach biegów i gąsienicach zdobycznych czołgów, a także przeglądali zdobyte dokumenty[1]. Mimo, że nie posiadano wyczerpujących danych, kolejne numery sprzętu i regularności w kodach umożliwiły Aliantom estymację produkcji czołgów niemieckich. Powojenne badania wykazały, że analiza numerów seryjnych dała bardziej dokładne oceny od zawyżonych estymat uzyskanych konwencjonalnymi metodami analizy (tab. 1)[2]. Zobacz artykuł Richarda Ruggles'a i Henry'iego Brodie'go [44], opisujący szczegółowo historię analizy numerów seryjnych, której użyto do oceny niemieckiej produkcji broni w czasie II wojny światowej.

PNC

Uproszczenie historycznego kontekstu estymacji produkcji czołgów na podstawie analizy numerów seryjnych [44] było motywem do sformułowania problemu w postaci dogodnej dla dyskusji o PNC w publikacji [21].

Sformułowanie problemu. W czasie II wojny światowej armia III Rzeszy miała na stanie n czołgów. Każdy czołg miał unikatowy numer seryjny ze zbioru liczb $\{1, \ldots, n\}$. Jako Alianci nie znamy n, ale zdobyliśmy próbkę k wrażych (wrogich – przyp. red.) czołgów (oczywiście jest to losowanie bezzwrotne), z których każdy miał swój numer seryjny z uporządkowanego zbioru $\{s_1, \ldots, s_k\}$.

1. Na przykład zdobyte dokumenty ze składów części zamiennych do czołgów zawierały listy numerów seryjnych podwozi i silników zreperowanych czołgów, a dokumenty ze sztabów dywizji zawierały numery seryjne czołgów danej jednostki.

2. Na przykład skrzynie biegów miały wybite numery seryjne jednym ciągiem. Numery podwozi natomiast były rozbite na bloki z przerwami między nimi. Pozwalało to odróżnić blok opisujący model i typ od bloku z numerem seryjnym.

*Dane oryginału artykułu: Simon, C.M. A Bayesian Treatment of the German Tank Problem. *Math Intelligencer* **46**, 117–127 (2024). https://doi.org/10.1007/s00283-023-10274-6

**ORCID: 0000-0002-8181-9178

Zakładając, że prawdopodobieństwo zdobycia każdego czołgu jest takie samo, a n jest nieznane, naszym celem jest estymacja n na podstawie danych $\{s_1, \ldots, s_k\}$.

W roku 1942 Alan Turing i Andrew Gleason, siedząc w zatłoczonej restauracji w Waszyngtonie, dyskutowali o pewnym wariancie PNC, a mianowicie: jak ocenić liczbę taksówek (słynnych ang. *yellow cabs* – przyp. tłum.) w mieście na podstawie zaobserwowanych losowych numerów przejeżdżających taksówek [13, 24]. Nawet dziś, ze względu na ciekawe tło historyczne, jest to ciągle interesującym tematem rozmowy przy stole i może posłużyć jako intelektualnie ciekawy, intrygujący i zabawny problem ilustrujący zagadnienia kombinatoryki i teorii estymacji podczas zajęć szkolnych [3, 15, 27, 33].

Tab. 1. Miesięczna produkcja czołgów niemieckich [44]

miesiąc	estymaty		dane niemieckie
	tradycyjne oceny angielskiego i amerykańskiego wywiadu	analiza numerów seryjnych	
czerwiec 1942	1000	169	122
lipiec 1942	1550	244	271
sierpień 1942	1550	327	342

Szacowanie niepewności oceny. Każda estymacja liczebności czołgów n na podstawie danych $\{s_1, \ldots, s_k\}$ jest obarczona niepewnością, gdyż nie zdobyliśmy (zapewne!) wszystkich czołgów (tzn. prawdopodobnie $k \neq n$). Znalezienie rozrzutu (błędu) estymacji jest ważne, gdyż na podstawie estymat podejmuje się istotne decyzje wojskowe.

Nasz wkład. W niniejszym dydaktycznym artykule wprowadzamy zarys bayesowskiego podejścia do PNC, którego rozwiązanie określa prawdopodobieństwo każdej możliwej liczby czołgów, co pozwala na ilościową ocenę niepewności estymacji. Daje to także możliwość wprowadzania do obliczeń wcześniejszej wiedzy lub przekonania o liczbie czołgów.

Przegląd wcześniejszych prac nad PNC

Wnioskowanie częstościowe (ang. *frequentist approach*). Kim Border [7] nazywa PNC „przedziwnym przypadkiem" estymowania za pomocą wnioskowania częstościowego. Zasada największej wiarygodności estymuje liczbę czołgów n jako maksymalny numer seryjny wśród k zdobycznych niemieckich czołgów $m^{(k)} = \max_{i\in\{1,\ldots,k\}} s_i$. Jest to estymator obciążony, gdyż na pewno $m^{(k)} \leqslant n$. Leo Goodman [21, 22] wyprowadził nieobciążony estymator o najmniejszej wariancji liczby czołgów

$$\hat{n} = m^{(k)} + \left(\frac{m^{(k)}}{k} - 1\right). \tag{1}$$

Aby wyrobić sobie intuicję co do wzoru (1) zauważmy, że $\hat{n}$ musi być większe od lub równe $m^{(k)}$, a jeśli rozważymy duże (lub małe) odstępy pomiędzy numerami seryjnymi $(s_1, \ldots, s_k)$, włączając w to odstęp pomiędzy zerem a najmniejszym numerem, to n jest zapewne dużo większe (lub nie) od $m^{(k)}$. Estymator n ze wzoru (1) wskazuje, o ile n powinno być większe od $m^{(k)}$ w zależności od obserwowanych odstępów; $m^{(k)}/k - 1$ to średni odstęp między numerami seryjnymi. Goodman [21] także wyprowadził wzór na dwustronny $(1 - a)$ przedział ufności $m^{(k)} \leqslant n \leqslant x$, gdzie x to największa liczba całkowita spełniająca $(m^{(k)} - 1)_k/(x)_k \geqslant a$ (oznaczenie $(n)_k$ dla malejącej silni zdefiniowano w [5]).

Zastosowania w dydaktyce. Julian Champkin [23] podkreśla, że użycie statystyki do estymacji niemieckiej produkcji czołgów w czasie II wojny światowej to był „wielki moment dla statystyki". Roger Johnson [27] podaje i ocenia wiele intuicyjnie jasnych punktowych estymatorów liczby czołgów. Richard Schaeffer i in. [45] proponują doświadczalne oszacowanie przez uczniów liczby ponumerowanych żetonów w worku poprzez wylosowanie kilku z nich i zanotowanie ich numerów jako model dla PNC. Arthur Berg [3] organizuje w klasie konkurs na znalezienie najlepszego estymatora liczby ludności w mieście na podstawie losowej próbki. George Clark i in. [10] badają użycie symulacji zdobywania wrażych czołgów i regresji liniowej do wykrycia estymatora ze wzoru (1).

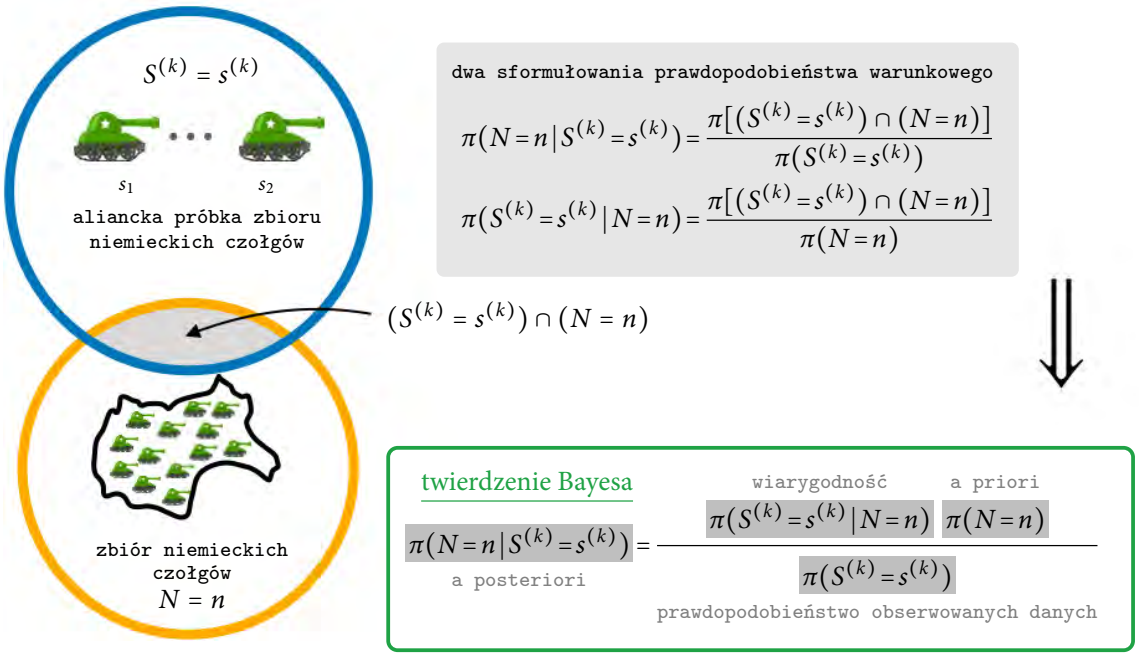
Podejście bayesowskie. Blisko związane z naszymi dydaktycznymi badaniami bayesowskiego podejścia do PNC są prace opublikowane przez innych badaczy. Harry Roberts [41], Michael Hoele i Leonard Held [25], Wolfgang Von der Linden, Volker Dose, Udo Von Toussaint [49], Simona Cocco, Remi Monasson, Francesco Zamponi [11] podejmują się bayesowskiej analizy PNC i wyprowadzają wzór analityczny na średnią i wariancję dla rozkładu prawdopodobieństwa *a posteriori* liczby czołgów w przypadku, gdy rozkład prawdopodobieństwa *a priori* jest płaski. Mark Andrews [1] daje zarys podejścia bayesowskiego do PNC na swoim blogu, gdzie także znajduje się program napisany w języku programowania R (służącym do obliczeń statystycznych i wizualizacji danych – przyp. red.). William Rosenberg i John Deely [43] dają zarys empirycznego podejścia dla estymacji liczby koni na wyścigach na podstawie próbki numerów kilku koni (funkcja wiarygodności jest równoważna tej dla PNC). Arthur Berg i Naur Hawila [4] posługują się wnioskowaniem bayesowskim w podobnym problemie taksówek.

Warianty i uogólnienia problemu. Goodman [21, 22] oraz Clark, Gonye i Miller [10] rozważają wariant PNC,

w którym początkowy numer seryjny jest nieznany; innymi słowy n czołgów jest znakowane numerami $\{b+1,\ldots,n+b\}$, ale b oraz n są nieznane. Lee i Miller uogólniają problem do sytuacji, gdy numery seryjne pochodzą z ciągłego zbioru liczb, a nawet gdy te numery są wybrane z kuli lub hipersześcianu w przestrzeni dwu lub więcej wymiarowej.

Przegląd bayesowskiego podejścia do PNC

Z bayesowskiego punktu widzenia [6, 15, 46] traktujemy (nieznaną) liczbę całkowitą czołgów jako dyskretną zmienną losową N, aby modelować stan naszej niepewności. Rozkład prawdopodobieństwa N daje nam prawdopodobieństwo dla każdej możliwej liczby czołgów n. To prawdopodobieństwo jest miarą naszej wiary, być może skojarzonej z jakąś uprzednią wiedzą, że ta liczba wynosi n [20]. Rozkład prawdopodobieństwa N opisuje naszą niepewność. Zaobserwowane numery seryjne $(s_1,\ldots,s_k)$ zawierają informację o liczebie czołgów. Tak więc rozkład prawdopodobieństwa N zmienia się, kiedy weźmiemy pod uwagę dane pomiarowe $(s_1,\ldots,s_k)$. Znaczy to, że N ma rozkład prawdopodobieństwa przed pomiarem (*a priori*) i po pomiarze (*a posteriori*).



Rys. 1. Zastosowanie twierdzenia Bayesa do PNC. Diagram Eulera [32, 37] przedstawia dwie konfiguracje $S^{(k)} = s^{(k)}$ i $N = n$ za pomocą kół. Powierzchnia każdego koła jest proporcjonalna do prawdopodobieństwa konfiguracji, a powierzchnia przecięcia – do jednoczesnego wystąpienia obu konfiguracji $(S^{(k)} = s^{(k)}) \cap (N = n)$. Diagram Eulera ilustruje dwie konfiguracje jako przecinające się względne prawdopodobieństwa, co implikuje twierdzenie Bayesa [30].

Oto trzy elementy podejścia bayesowskiego do problemu PNC:

1) rozkład prawdopodobieństwa *a priori* N uwzględniający naszą wiedzę lub przekonanie o wartości n przed uzyskaniem próbki numerów seryjnych,
2) dane, czyli próbka numerów seryjnych $(s_1,\ldots,s_k)$ traktowanych jako wynik losowania zmiennych losowych $(S_1,\ldots,S_k)$ przy założeniu, że wybór numerów jest czysto stochastyczny (losowy),
3) wiarygodność wyrażająca prawdopodobieństwo obserwowanych danych $(S_1,\ldots,S_k) = (s_1,\ldots,s_k)$ dla każdej wartości $N = n$, zależnie od przyjętego statystycznego modelu przechwytywania czołgów.

Wynikiem baysowskiego podejścia do problemu PNC jest rozkład prawdopodobieństwa *a posteriori* na podstawie danych $(s_1,\ldots,s_k)$. Rozkład prawdopodobieństwa *a posteriori* można uważać za aktualizację rozkładu *a priori* w świetle nowych danych (rys. 1). Rozkład prawdopodobieństwa *a posteriori* dla N przypisuje każdej liczbie czołgów n prawdopodobieństwo będące kompromisem między wiarygodnością (odwołującą się do modelu pozyskiwania zdobycznych czołgów i obserwowanych danych $(s_1,\ldots,s_k)$ oraz hipotezy, że liczba czołgów wynosi n), a rozkładem prawdopodobieństwa *a priori* N, które wynika z naszego przekonania o wartości n przed uzyskaniem próbki numerów seryjnych $(s_1,\ldots,s_k)$ [46]. Rozkład prawdopodobieństwa *a posteriori* jest wyjściowym wynikiem dla dalszej analizy i obrazuje, jak rozrzucone są prawdopodobieństwa w N. Możemy uzyskać sumaryczny wynik analizy podając np. medianę i mały zbiór liczb, dla któ-

rych prawdopodobieństwo jest największe, czyli wiarygodny zbiór, w którym zapewne znajduje się rzeczywista liczba czołgów. Dalej, na podstawie rozkładu *a posteriori* możemy odpowiedzieć na pytanie : jakie jest prawdopodobieństwo, że n będzie większe od pewnej progowej wielkości n', która mogłoby zmienić wojskową strategię.

Tab. 2. Lista parametrów/zmiennych

parametry/zmienne	$\in$	opis
n	$\mathbb{N}_{\geqslant 0}$	liczba czołgów
k	$\mathbb{N}_{>0}$	liczba zdobytych czołgów
s_i	$\mathbb{N}_{>0}$	numer seryjny i-tego czołgu
$s^{(k)}$	$\mathbb{N}_{>0}^{k}$	wektor numerów seryjnych k zdobycznych czołgów
$m^{(k)}$	$\mathbb{N}_{>0}$	maksymalny numer seryjny wśród k zdobycznych czołgów

Bayesowskie podejście do PNC

Teraz zajmiemy się szczegółami podejścia bayesowskiego do PNC i zilustrujemy to przykładem (użyte dalej symbole podane są w tab. 2). Używamy dużych liter dla zmiennych losowych a małych – dla ich konkretnych realizacji. Używamy funkcji charakterystycznej związanej ze zbiorem A:

$$\mathcal{I}_A(x) = \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases} \tag{2}$$

Rozkład prawdopodobieństwa *a priori*. Tworzymy rozkład prawdopodobieństwa *a priori*, będący kombinacją naszych subiektywnych ocen i/lub obiektywnej wcześniejszej wiedzy o $\pi_{\text{a priori}}(N = n)$ przed uzyskaniem próbki numerów seryjnych $(s_1, \ldots, s_k)$. Rozkład prawdopodobieństwa *a priori* N, który postulujemy, zależy od okoliczności. Jeśli nie mamy żadnej informacji a *priori* o liczbie czołgów, to możemy przyjąć zasadę niewyróżniana żadnych parametrów i użyć rozmytego rozkładu *a priori* np. rozkładu płaskiego. W przeciwnym wypadku, kiedy mamy zgrubne pojęcie o liczbie czołgów z innych źródeł lub analiz, możemy opisać to rozkładem zawierającym więcej informacji, np. rozkładem skoncentrowanym wokół naszej najlepszej uprzedniej estymacji. Zgodnie z definicją rozmyte rozkłady zawierają większą nieoznaczoność, mierzoną np. entropią [35]. (We współczesnej statystyce, w sytuacji, gdy brakuje nam informacji o danym parametrze, wybieramy rozkład o największej entropii informacyjnej Shannona; w rozważanym przypadku jest to rozkład płaski [46] – przyp. tłum.)

Jeśli teraz myślimy o rozkładzie prawdopodobieństwa *a posteriori*, który balansuje między rozkładem *a priori* a wiarygodnością (wykorzystującą dane), to rozkład *a priori*, niosący więcej informacji niż rozkład rozmyty, będzie miał większy wpływ na rozkład *a posteriori* niż rozkład rozmyty, który „pozwala danym mówić samodzielnie" [15]. Zasadniczo, kiedy liczba zdobycznych czołgów k rośnie, to spodziewamy się, że rozkład *a priori* będzie miał coraz mniejszy wpływ na rozkład *a posteriori* [15], jako że dane „zdominują" rozkład *a priori*.

Dane, generowanie danych i funkcja wiarygodności

Dane. Dane, którymi dysponujemy w PNC to wektor

$$s^{(k)} = (s_1, \ldots, s_k) \tag{3}$$

numerów seryjnych wybitych na k zdobycznych czołgach. Dane te traktujemy jako realizację wektora rozkładów losowych $S^{(k)} = (S_1, \ldots, S_k)$. W tym miejscu zakładamy, że kolejność pojmanych czołgów ma znaczenie.

Proces generowania danych. Stochastyczne generowanie danych polega na kolejnym, bezzwrotnym zdobywaniu k czołgów ze zbioru n czołgów, a następnie zapisywaniu ich numerów do $(s_1, \ldots, s_k)$. Zakładamy, że w każdym kroku dowolny czołg ma takie samo prawdopodobieństwo bycia zdobytym. Z matematycznego punku widzenia to jest losowanie bezzwrotne k liczb całkowitych z rozkładu płaskiego ze zbioru $\{1, \ldots, n\}$.

Funkcja wiarygodności. Funkcja wiarygodności daje prawdopodobieństwo obserwacji danych $(S_1, \ldots, S_k) = (s_1, \ldots, s_k)$ przy założeniu, że liczba czołgów $N = n$. Każdy wynik $s^{(k)}$ w przestrzeni próbek $\Omega_n^{(k)}$ jest tak samo prawdopodobny, gdzie

$$\Omega_n^{(k)} = \{(s_1, \ldots, s_k)_{\neq} : s_i \in \{1, \ldots, n\} \text{ dla wszystkich } i \in \{1, \ldots, k\}\}, \tag{4}$$

a zapis $(\cdots)_{\neq}$ oznacza, że elementy wektora $(\cdots)$ są różne. Liczba wyników $|\Omega_n^{(k)}|$ w przestrzeni próbek, to liczba różnych uporządkowań k unikatowych liczb ze zbioru $\{1, \ldots, n\}$ dana wzorem na malejącą silnię

$$(n)_k = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}. \tag{5}$$

W czasie generowania danych prawdopodobieństwo zaobserwowania $S^{(k)} = s^{(k)}$ przy założeniu, że rozkład liczby czołgów $N = n$ jest płaski

$$\pi_{\text{wiarygodność}}(S^{(k)} = s^{(k)} \mid N = n) = \frac{1}{(n)_k} \mathcal{I}_{\Omega_n^{(k)}}(s^{(k)}). \tag{6}$$

Interpretacja. Wiarygodność dana jest wzorem, z którego wynika, jak mocne jest poparcie dla naszej wiedzy wynikającej z analizy numerów seryjnych k zdobycznych

czołgów w $s^{(k)}$ w porównaniu z naszym probabilistycznym modelem zdobywania wrażych czołgów, przy hipotezie, że liczba czołgów to n [46]. Traktujemy wyrażenie $\pi_{\text{wiarygodność}}(S^{(k)} = s^{(k)} \min N = n)$ jako funkcję n, gdyż mamy tylko do dyspozycji dane $s^{(k)}$, a nie n.

Wiarygodność jako sekwencja zdarzeń. Możemy wyprowadzić wzór (6) w inny sposób. Rozważmy ciąg zdarzeń $S_1 = s_1, S_2 = s_2, \ldots, S_k = s_k$. Prawdopodobieństwo wystąpienia danego numeru na zdobytym czołgu, biorąc pod uwagę liczbę czołgów i numery seryjne uprzednio zdobytych czołgów, wynika z rozkładu płaskiego

$$\pi(S_i = s_i \mid N = n,\ S_1 = s_1, \ldots, S_{i-1} = s_{i-1}) = \frac{1}{n - i + 1}\mathcal{I}_{\{1,\ldots,n\}\setminus\{s_1,\ldots,s_{i-1}\}}(s_i), \quad (7)$$

skoro mamy do dyspozycji $n - (i - 1)$ czołgów do losowego wybrania z rozkładu płaskiego. Z reguły mnożenia niezależnych prawdopodobieństw [29] sumaryczne prawdopodobieństwo to

$$\pi_{\text{wiarygodność}}(S_1 = s_1, \ldots, S_k = s_k \mid N = n) = \prod_{i=1}^{k} \pi(S_i = s_i \mid N = n,\ S_1 = s_1, \ldots, S_{i-1} = s_{i-1}), \quad (8)$$

co daje wzór (6) po wykorzystaniu własności funkcji charakterystycznej.

Funkcja wiarygodności w zależności od maksymalnego obserwowanego numeru seryjnego. Przekonamy się tu, że tylko dwie niezależne cechy danych $(s_1, \ldots, s_k)$ podają informację o liczbie czołgów N: długość próbki k oraz obserwowany maksymalny numer seryjny

$$m^{(k)} = \max_{i\in\{1,\ldots,k\}} s_i. \quad (9)$$

Tak więc szukamy tu innej funkcji wiarygodności, tj. prawdopodobieństwa $\pi_{\text{wiarygodność}}(M^{(k)} = m^{(k)} \mid N = n)$ zaobserwowania maksymalnego numeru seryjnego $m^{(k)}$ dla zadanej liczby $N = n$. Każdy wynik $s^{(k)} \in \Omega_n^{(k)}$ jest tak samo prawdopodobny, więc $\pi_{\text{wiarygodność}}(M^{(k)} = m^{(k)} \mid N = n)$ jest ułamkiem przestrzeni próbek $\Omega_n^{(k)}$, w której maksymalny numer seryjny to $m^{(k)}$. Aby obliczyć liczbę wyników $s^{(k)} \in \Omega_n^{(k)}$, gdzie maksymalny numer seryjny to $m^{(k)}$, rozważmy taką sytuację, że jeden z k zdobycznych czołgów ma numer seryjny $m^{(k)}$, a pozostałe $k - 1$ mają numery ze zbioru $\{1, \ldots, m^{(k)} - 1\}$. Dla każdej z k możliwych pozycji maksymalnego numeru seryjnego w wektorze $s^{(k)}$ mamy $(m^{(k)} - 1)_{k-1}$ różnych wyników określających resztę $k - 1$ numerów. Tak więc

$$\pi_{\text{wiarygodność}}(M^{(k)} = m^{(k)} \mid N = n) = \frac{k(m^{(k)} - 1)_{k-1}}{(n_k)}\mathcal{I}_{\{k,\ldots,n\}}(m^{(k)}). \quad (10)$$

Rozkład prawdopodobieństwa *a posteriori*

Gęstość prawdopodobieństwa *a posteriori* jako funkcja N przypisuje prawdopodobieństwo każdej możliwej liczby czołgów n uwzględniając dane $(s_1, \ldots, s_k)$ zgodnie z funkcją wiarygodności (6) i naszymi uprzednimi przekonaniami i wiedzą opisane przez $\pi_{\text{a priori}}(N = n)$.

Rozkład a *posteriori* jest warunkowym rozkładem wynikającym z wiarygodności i z rozkładu prawdopodobieństwa *a priori* zgodnie z twierdzeniem Bayesa [30] (rys. 1)

$$\pi_{\text{a posteriori}}(N = n \mid S^{(k)} = s^{(k)}) = \frac{\pi_{\text{wiarygodność}}(S^{(k)} = s^{(k)} \mid N = n)\pi_{\text{a priori}}(N = n)}{\pi_{\text{evidence}}(S^{(k)} = s^{(k)})}. \quad (11)$$

Mianownik, czyli prawdopodobieństwo obserwowanych danych (ang. *evidence*) $s^{(k)}$ [30] jest dane wzorem

$$\pi_{\text{evidence}}(S^{(k)} = s^{(k)}) = \sum_{n'=0}^{\infty} \pi_{\text{wiarygodność}}(S^{(k)} = s^{(k)} \mid N = n')\pi_{\text{a priori}}(N = n'). \quad (12)$$

Rozkład $\pi_{\text{a posteriori}}(N = n \mid S^{(k)} = s^{(k)})$ traktujemy jako rozkład prawdopodobieństwa dla N, skoro w praktyce mamy zawsze do dyspozycji $s^{(k)}$. Niezależne od n $\pi_{\text{evidence}}(S^{(k)} = s^{(k)})$, jest tylko czynnikiem normalizacyjnym w mianowniku (11). Wzór (11) interpretujemy następująco: gęstość prawdopodobieństwa *a priori* jest aktualizowana w świetle obserwowanych danych $(s_1, \ldots, s_k)$ dając w wyniku rozkład prawdopodobieństwa *a posteriori* dla N. Prawdopodobieństwo *a posteriori* dla $N = n$ jest proporcjonalne do iloczynu wiarygodności i rozkładu *a priori* dla $N = n$, co stanowi kompromis pomiędzy wiedzą *a priori* a wiarygodnością. Możemy uprościć rozkład prawdopodobieństwa *a posteriori* (11) podstawiając (6) i ograniczając sumę w (12) do liczby i czołgów, dla których wiarygodność jest niezerowa. Zauważmy, że tylko dwie dane ze zbioru danych $(s_1, \ldots, s_k)$ pojawiają się we wzorze i są to długość danych k i maksymalny numer seryjny $m^{(k)}$

$$\begin{aligned}\pi_{\text{a posteriori}}(N = n \mid S^{(k)} = s^{(k)}) &= \pi_{\text{a posteriori}}(N = n \mid M^{(k)} = m^{(k)}) \\ &= \frac{(n)_k^{-1}\pi_{\text{a priori}}(N = n)}{\sum_{n'=m^{(k)}}^{\infty}(n')_k^{-1}\pi_{\text{a priori}}(N = n')}\mathcal{I}_{\{m^{(k)}, m^{(k)}+1, \ldots\}}(n).\end{aligned} \quad (13)$$

Zauważmy, że możemy także otrzymać (13) posługując się (10).

Interpretacja. Gęstość prawdopodobieństwa *a posteriori* jako funkcja N (13) określa prawdopodobieństwo liczby czołgów n, przy wzięciu pod uwagę numerów seryjnych $(s_1, \ldots, s_k)$ zaobserwowanych na zdobycznych czołgach, nasz model probabilistyczny procesu zdobywania wrażych czołgów i naszą wiedzę *a priori,* czyli kombinację naszych subiektywnych ocen i/lub obiektywnej uprzedniej wiedzy wyrażoną rozkładem *a priori* w funkcji N. Rozrzut (mierzony np. entropią informacyjną) rozkładu *a posteriori* odzwierciedla epistemiczną (redukowalną przez więcej danych) [17, 47] niepewność naszej wiedzy o liczbie czołgów.

Uwaga o „niepewności". Źródłem niepewności zawartej w rozkładzie *a posteriori* jest brak pełnych danych: nie udało się zdobyć wszystkich wrażych czołgów, aby mieć pewność, co do ich liczby[3]. W praktyce dodatkowym źródłem niepewności jest możliwa nieadekwatność modelu zdobywania wrażych czołgów (założenie próbkowania z płaskiego rozkładu) we wzorze (6). Innymi słowy, w opisie procesu zdobywania wrażych czołgów mogłoby zachodzić odchylenie od modelu próbkowania z płaskiego rozkładu. W naszej analizie pomijamy taką możliwość.

Streszczenie opisu rozkładu a *posteriori*. Możemy streścić naszą wiedzę zawartą w rozkładzie *a posteriori* podając punktową estymatę i mały zbiór liczb, dla których prawdopodobieństwo jest największe – wiarygodny zbiór, w którym zapewne znajduje się rzeczywista liczba czołgów[4]. Pożyteczną punktową estymatą jest mediana rozkładu *a posteriori*; z definicji jest to wartość, dla której prawdopodobieństwo, że n jest większe lub równe medianie, wynosi $\frac{1}{2}$. Właściwy wiarygodny podzbiór liczby wrażych czołgów to podzbiór a [26].

$$\mathcal{H}_a = \{n' : \pi_{\text{a prosteriori}}(N = n' \mid M^{(k)} = m^{(k)}) \geqslant \pi_a\}, \tag{14}$$

gdzie π_a to największe n spełniające

$$\pi_{\text{a prosteriori}}(N \in \mathcal{H}_a \mid M^{(k)} = m^{(k)}) \geqslant 1 - a. \tag{15}$$

Innymi słowy, a, będący podzbiorem $\mathcal{H}_a$, to najmniejszy zbiór, który zawiera część $(1 - a)$ rozkładu *a posteriori* dla takich n, że każde z nich występuje z większym prawdopodobieństwem niż dowolne n spoza tego zbioru.

Wnioskowanie na podstawie rozkładu *a posteriori*. Posługując się rozkładem *a posteriori* możemy wyznaczyć dowolny zbiór n spełniających zadane warunki; np. prawdopodobieństwo tego, że liczba wrogich czołgów jest większa od jakiegoś n' wynosi

$$\pi_{\text{a prosteriori}}(N = n' \mid M^{(k)} = m^{(k)}) = \sum_{n=n'+1}^{\infty} \pi_{\text{a prosteriori}}(N = n \mid M^{(k)} = m^{(k)}). \tag{16}$$

Przykład

Poniższym przykładem ilustrujemy podejście bayesowskie do PNC.

Rozkład prawdopodobieństwa *a priori* w zależności od N. Załóżmy, że mamy kres górny $n_{\max}$ możliwej liczby wrażych czołgów na podstawie np. dostaw pewnych surowców koniecznych do produkcji czołgów, a poza tym żadnych innych informacji. Możemy wtedy wybrać rozmyty rozkład prawdopodobieństwa *a priori* – rozkład płaski

$$\pi_{\text{a priori}}(N = n) = \frac{1}{n_{\max} + 1} \mathcal{I}_{\{0,\ldots,n_{\max}\}}(n). \tag{17}$$

Ten rozkład prawdopodobieństwa *a priori* wyraża fakt, że przy braku jakichkolwiek danych $(s_1, \ldots, s_k)$ (tzn. żadnych numerów seryjnych ani nawet liczby k), sądzimy, że całkowita liczba wrażych czołgów N może przyjąć dowolną wartość ze zbioru $(0, \ldots, n_{\max})$ z takim samym prawdopodobieństwem. Dla ustalenia uwagi załóżmy, że $n_{\max} = 35$. Na rysunku 2(a) widzimy $\pi_{\text{a priori}}(N = n)$.

Dane $(s_1, \ldots, s_k)$ i funkcja wiarygodności. Załóżmy teraz, że zdobyliśmy k=3 czołgi (rys. 2(b)) z numerami seryjnymi $s^{(3)} = (15, 14, 3)$. Maksymalny numer seryjny to $m^{(3)} = 15$. Funkcja wiarygodności ze wzoru (10) $\pi_{\text{wiarygodność}}(M(3) = 15 \mid N = n)$ pokazana jest na rysunku 2(c). Zauważmy, że funkcja wiarygodności osiąga maksimum dla $n = m^{(3)} = 15$, a potem maleje jednostajnie.

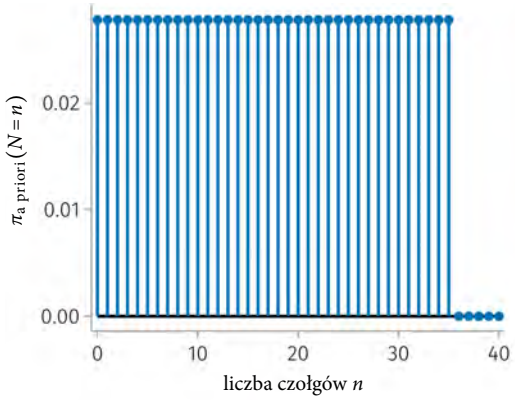
Rozkład prawdopodobieństwa *a posteriori* w zależności od N. Uwzględniając rozkład *a priori* (17), ze wzoru (13) otrzymujemy

$$\pi_{\text{a prosteriori}}(N = n \mid M^{(k)} = m^{(k)}) = \frac{(n)_k^{-1}}{\sum_{n'=m^{(k)}}^{n_{\max}}} \mathcal{I}_{\{m^{(k)}, m^{(k)}+1, \ldots, n_{\max}}(n). \tag{18}$$

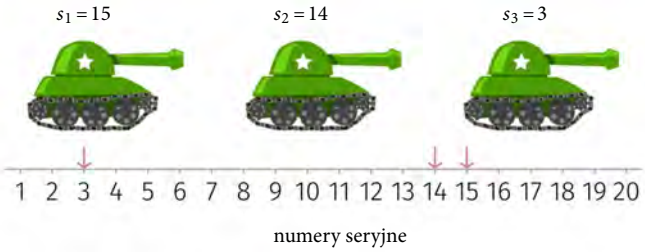
Na rysunku 2(d) pokazany jest rozkład prawdopodobieństwa *a posteriori* dla danych $s^{(3)}$ z rysunku 2(b) i rozkład prawdopodobieństwa *a priori* opisany wzorem (17) ($n_{\max} = 35$).

3. Na pewno $k < n$, jeżeli są luki pomiędzy obserwowanymi numerami seryjnymi $(s_1, \ldots, s_k)$. Nawet gdy nie ma luk w $(s_1, \ldots, s_k)$, to nie możemy być pewni, że zdobyliśmy wraży czołg o największym numerze seryjnym.

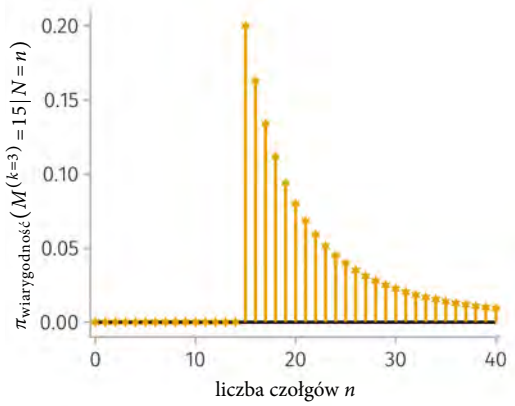
4. przy naszych założeniach o funkcji wiarygodności i rozkładzie *a priori*



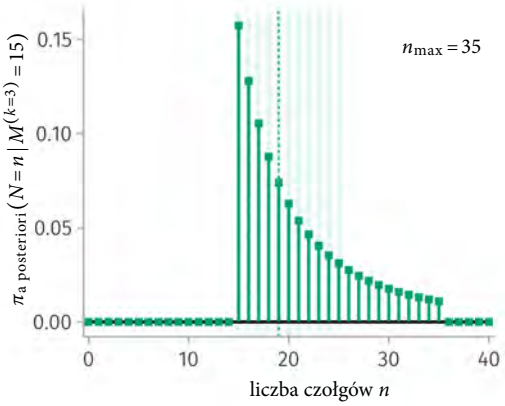
(a) funkcja liczby czołgów N *a priori*



(b) dane $s^{(k=3)}$

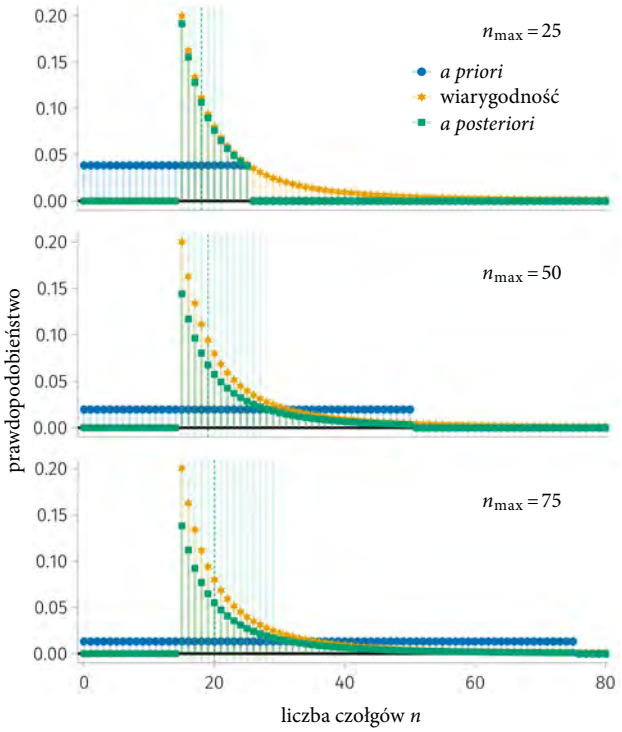


(c) funkcja wiarygodności



(d) funkcja liczby czołgów N *a posteriori*

Rys. 2. Podejście bayesowskie do PNC. (a) Rozkład prawdopodobieństwa *a priori*. (b) Dane $s^{(3)}$ z maksymalnym zaobserwowanym numerem seryjnym $m^{(3)} = 15$. (c) Funkcja wiarygodności dla danych $s^{(3)}$. (d) Rozkład prawdopodobieństwa *a posteriori*; $\mathcal{H}_{0.2}$ jest wyróżnione kolorem a mediana – pionową linią przerywaną.



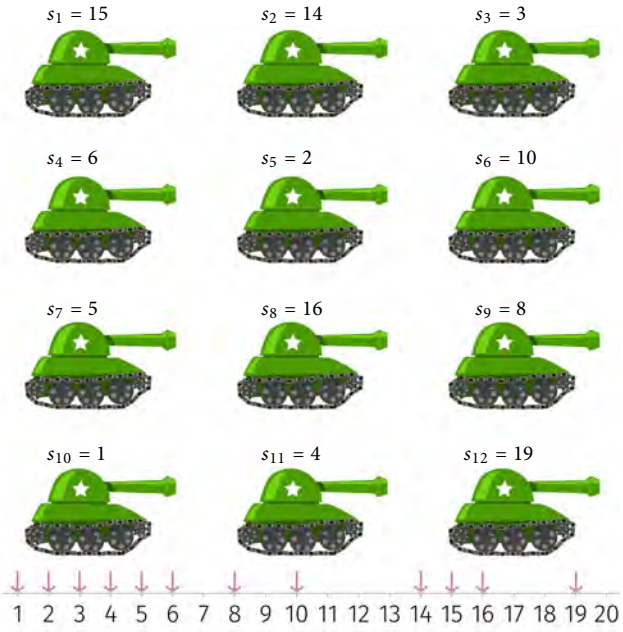
Rys. 3. Czułość rozkładu *a posteriori* na zmiany kresu górnego $n_{\max}$ użytego w rozkładzie *a priori*.

Podsumowanie opisu rozkładu *a posteriori*. Rozkład a posteriori jako funkcja N ma medianę 19 i wiarygodny podzbiór $\mathcal{H}_{0.2} = \{15, \ldots, 25\}$ wyróżniony kolorem na rysunku 2(d); dane te zostały wygenerowane komputerowo dla liczby czołgów $n = 20$, co uzasadnia skalę użytą na rysunku 2(b).

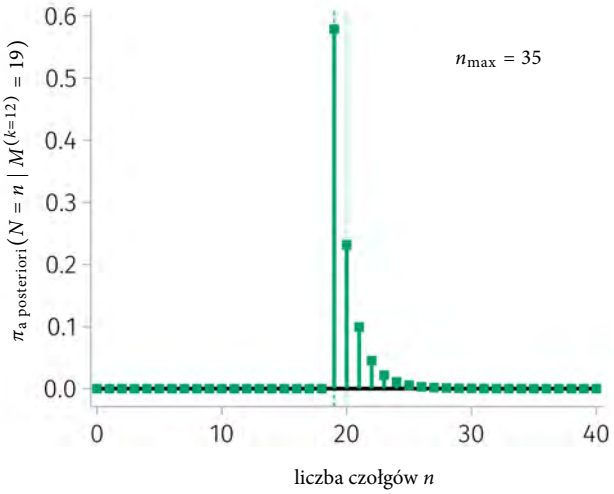
Wnioskowanie na podstawie rozkładu *a posteriori*. Załóżmy, że strategia Aliantów musiałaby ulec zmianie, gdyby liczba wrażych czołgów przekroczyła 30. Z rozkładu *a posteriori* (N) obliczamy, że $\pi_{\text{a posteriori}}(N > 30 \mid M^{(3)} = 15) \approx 0{,}066$.

Czułość rozkładu *a posteriori* na zmiany w rozkładzie *a priori*. Skoro przy konstruowaniu rozkładu *a priori* odgrywa rolę czynnik subiektywny, jest dobrą praktyką sprawdzić czułość rozkładu *a posteriori* na zmiany w rozkładzie *a priori* [46]. Na rysunku 3 widzimy, jak zmienia się rozkład *a posteriori*, gdy zmieniamy górny kres liczby wrażych czołgów $n_{\max}$, którego używamy we wzorze (17). Na przykład, gdy zmienimy na $n_{\max} = 75$, to wiarygodny podzbiór zwiększa się do $\mathcal{H}_{0.2} = \{15, \ldots, 29\}$.

Efekt zdobycia większej liczby wrogich czołgów. Załóżmy, że zdobyliśmy dodatkowych 9 czołgów i pow-



(a) zaktualizowane dane $s^{(k=12)}$



(b) zaktualizowany rozkład *a posteriori*

Rys. 4. Rozkład *a posteriori* po zdobyciu większej liczby wrażych czołgów. (a) Zdobyliśmy dodatkowych 9 czołgów. (b) Zaktualizowany rozkład *a posteriori* jako funkcja N.

tórzmy analizę bayesowską. Na rysunku 4 pokazany jest zaktualizowany rozkład *a posteriori*. Wiarygodny podzbiór $\mathcal{H}_{0.2}$ zmniejsza się znacznie do $\{19, 20\}$. Widać, jak więcej danych (większa liczba zdobycznych czołgów) ogólnie zmniejsza naszą niepewność oceny liczby wrażych czołgów.

Badania zachowania rozkładu *a posteriori* przy zmianie znanej liczby wrażych czołgów, za pomocą symulacji komputerowych

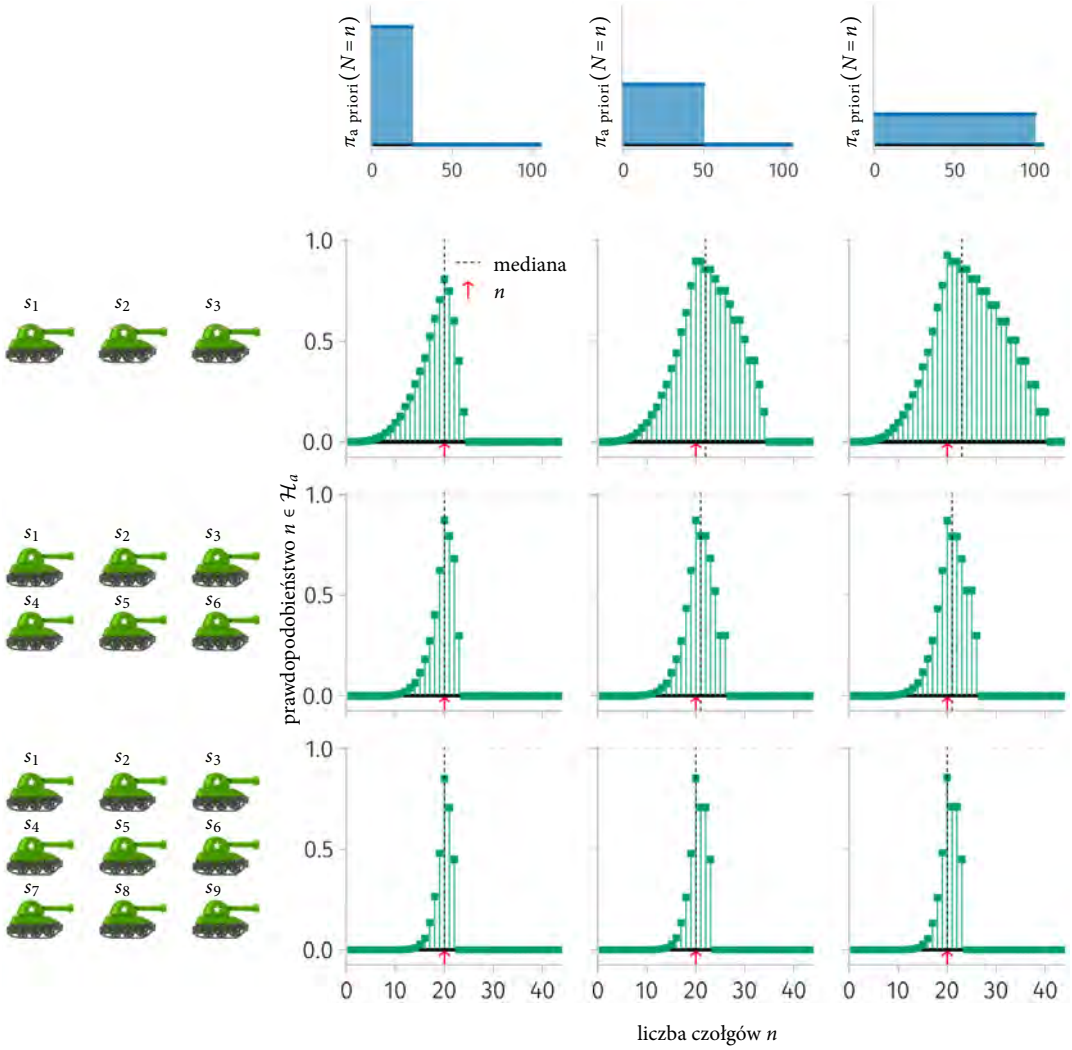
Badamy tu, jak średnio symulowane wyniki zdobywania wrogich czołgów, przy założeniu ustalonej ich liczy, pokazują zależność prawdopodobieństwa *a posteriori* liczby czołgów $\pi_{\text{a posteriori}}(N = n \mid S^{(k)} = s^{(k)})$ od liczby k zdobycznych czołgów i od maksimum $n_{\max}$ w postulowanym płaskim rozkładzie *a priori*.

Dla zadanego k i $n_{\max}$ wykonujemy 50 000 symulacji. W każdej wybieramy losowo k wrażych czołgów ze zbioru $n = 20$, obliczamy prawdopodobieństwa *a posteriori* liczby czołgów $\pi_{\text{a posteriori}}(N = n \mid S^{(k)} = s^{(k)})$; następnie obliczamy medianę i wiarygodny podzbiór $\mathcal{H}_a(a = 0.2)$. Na rysunku 5 widzimy w pierwszej kolumnie liczbę zdobycznych czołgów, w pierwszym wierszu rozkłady *a priori*; dalej mamy tabelę 3×3 wykresów: rozkładu *a posteriori*, mediany i wiarygodnego podzbioru dla $(k, n_{\max}) \in \{3.6, 9\} \times \{25, 50, 100\}$. W miarę wzrostu k, wiarygodny podzbiór $\mathcal{H}_a$ staje się mniej czuły na $n_{\max}$, gdyż dane zaczynają dominować nad rozkładem *a priori*. Zbiór $\mathcal{H}_a$ staje się coraz mniej liczny, jako że niepewność zmniejsza się wraz ze wzrostem liczebności próbki. Ze wzrostem $n_{\max}$, $\mathcal{H}_a$ zawiera coraz to większe wartości n (szacowane liczby wrogich czołgów). Mediana uzyskanych w symulacji median rozkładów *a posteriori* pokrywa się z prawdziwą liczbą czołgów = 20 dla $n_{\max} = 25$ lub $k = 9$. Dla wartości $k \in \{3, 6\}$ większe wartości $n_{\max}$ przesuwają medianę powyżej prawdziwej liczby czołgów.

Dyskusja

Błąd selekcji. Formułując podręcznikowo PNC przyjmujemy mocne założenie, które pozwala estymować liczbę wrażych czołgów na podstawie losowej próbki numerów seryjnych na nich wybitych, tj. postulujemy, że numery seryjne są wybrane ze zbioru liczb $(1, 2, \ldots)$ i że każda liczba jest równie prawdopodobna (rozkład płaski); Goodman pokazał w [22] test na tę hipotezę. Interesującą wariacją na temat PNC mogłoby być modelowanie innych rozkładów prawdopodobieństwa zaobserwowania danego numeru seryjnego. Na przykład moglibyśmy uwzględniać fakt, że starsze czołgi, z mniejszymi numerami seryjnymi, mogły być wprowadzone do walki wcześniej. To mogłoby spowodować, że starsze wraże czołgi byłyby trudniejsze do zdobycia niż te nowsze, jako że późniejsze fronty były mniej ufortyfikowane. Innym przykładem błędu selekcji mogłoby być to, że numery seryjne grupują się w grupy blisko leżących liczb.

PNC w innych kontekstach. Bayesowskie podejście użyte do analizy PNC można zastosować (być może z jakimiś modyfikacjami) w innych sytuacjach, gdy potrzebujemy oszacowania wielkości jakiegoś innego skończonego zbioru [9], np. liczby taksówek w mieście [19, 23], samochodów na torze wyścigowym [48], rachunków w banku [25], mebli zakupionych przez uniwersytet [22], operacji na lotnisku [34], rozpraw w sądzie [50] lub sprzętu elektronicznego wyprodukowanego w fabryce [2]. W ten sam sposób możemy szacować: liczbę tajnych dokumentów rządowych, które wypłynęły na jaw [18], czas potrzebny do zakończenia projektu [16],



Rys. 5. Średni wiarygodny podzbiór i mediana mediany rozkładu *a posteriori* przy stałej liczbie czołgów. Wiersze: różne liczby k zdobycznych czołgów (lewa część). Kolumny: różne założenia *a priori* o górnym kresie możliwej liczby czołgów $n_{\max}$ dla rozkładu płaskiego (góra). Dla każdej pary $(k, n_{\max})$ wykres pokazuje prawdopodobieństwo *a posteriori* liczby czołgów. Pionowe linie przerywane wskazują medianę, a czerwona strzałka – rzeczywistą liczbę równą 20.

okres opisu ekstremalnych zjawisk w przeszłości takich jak powodzie [39], długość krótkich tandemowych powtórzeń allelu (allel to jedna z wersji genu – przyp. red.) [51], rozmiar sieci społecznych w Internecie [28], czas życia kwiatu [38] lub czas trwania jakiegoś gatunku [42]. Dodatkowo w badaniach ekologicznych obrączkowanie i odławianie pozwala na ocenę liczebności zwierząt [8, 36]. Wszystko to jest spokrewnione z PNC.

Praktyka wybijania kolejnych numerów na sprzęcie wojskowym. W Niemczech wprowadzono praktykę oznaczania sprzętu wojskowego numerami seryjnymi i kodami pozwalającymi zidentyfikować producenta. Jednak użycie kolejnych numerów seryjnych zostało wykorzystane przez Aliantów do szacowania tempa produkcji uzbrojenia w III Rzeszy. Aby zmniejszyć zagrożenie szacowania produkcji analizą numerów seryjnych, przy zachowaniu zalet możliwości identyfikacji producenta, można zastosować szyfrowanie [14] lub konfundowanie przeciwnika metodą zwaną z ang. *chaffing* [40].

Podziękowania

Niniesza praca otrzymała wsparcie United States Department of Homeland Security Countering Weapons of Mass Destruction under CWMD Academic Research Initiative Cooperative Agreement #21CWDARI00043, które nie stanowi poparcia rządowego ani *implicite* ani jawnie wyrażonego.

Dziękuję Berhardowi Konradowi i Edwardowi Celarierowi za szczegółowe uwagi o manuskrypcie. Moi studenci: Gbenga Fabusola, Adrian Henle i Paul Morris wnieśli swoje uwagi na temat wstępu, za co im dziękuję. Anonimowemu recenzentowi należy się podziękowanie za sugestię napisania rozdziału *Badania zachowania rozkładu a posteriori przy zmianie znanej liczby wrażych czołgów, za pomocą symulacji komputerowych*,

a Paulowi Campbellowi za informację o przykładzie estymacji liczby iPhonów.

Dostęp do danych i kodu źródłowego

Program w języku programowania Julia [5] pozwala odtworzyć wszystkie nasze wyniki (z grafiką wykonaną Makie.jl [12]) i jest dostępny na Github https://www.github.com/SimonEnsemble/the_German_tank_problem

Open Access

Bibliografia

[1] Mark Andrews. German tank problem: a Bayesian analysis. Available at https://www.mjandrews.org/blog/germantank. Accessed 2022-12-03.

[2] Charles Arthur. Why iPhones are just like German tanks. Available at https://www.theguardian.com/technology/blog/2008/oct/08/iphone.apple, 2008.

[3] Arthur Berg. Bayesian modeling competitions for the classroom. *Revista Colombiana de Estadıstica* 44:2 (2021), 243–252.

[4] Arthur Berg, Nour Hawila. Introducing Bayesian inference with the taxicab problem. In *Proceedings of the Tenth Australian Conference on Teaching Statistics*, ss. 55–60, 2021.

[5] Jeff Bezanson, Stefan Karpinski, Viral B. Shah, Alan Edelman. Julia: A fast dynamic language for technical computing. arXiv:1209.5145, 2012.

[6] William M. Bolstad, James M. Curran. *Introduction to Bayesian Statistics*. John Wiley & Sons, 2016.

[7] Kim C. Border. Lecture 18: Estimation. Available at https://healy.econ.ohio-state.edu/kcb/Ma103/ (2021 version), 2017.

[8] Anne Chao. An overview of closed capture–recapture models. *Journal of Agricultural, Biological, and Environmen- tal Statistics* 6:2 (2001), 158–175.

[9] Si Cheng, Daniel J. Eck, Forrest W. Crawford. Estimating the size of a hidden finite set: large-sample behavior of estimators. *Statistics Surveys* 14 (2020), 1–31.

[10] George Clark, Alex Gonye, Steven J. Miller. Lessons from the German tank problem. *Mathematical Intelligencer* 43:4 (2021), 19–28.

[11] Simona Cocco, Rémi Monasson, Francesco Zamponi. *From Statistical Physics to Data-Driven Modelling, with Applications to Quantitative Biology*. Oxford University Press, 2022.

[12] Simon Danisch, Julius Krumbiegel. Makie.jl: Flexible high-performance data visualization for Julia. *Journal of Open Source Software* 6:65 (2021), 3349.

[13] Peter Donovan. Alan Turing, Marshall Hall, the alignment of WW2 Japanese naval intercepts. *Notices of the AMS* 61:3, 2014.

[14] Hans Delfs, Helmut Knebl, Helmut Knebl. *Introduction to Cryptography*, volume 2. Springer, 2002.

[15] Allen B. Downey. Think Bayes 2. Available at https://allendowney.github.io/ThinkBayes2/index.html, 2021.

[16] Thomas M. Fehlmann, Eberhard Kranich. A new approach for continuously monitoring project deadlines in software development. In *Proceedings of the 27th International Workshop on Software Measurement and 12th Interna- tional Conference on Software Process and Product Measure- ment*, ss. 161–169, 2017.

[17] Craig R. Fox, Gülden Ülkümen. Distinguishing two dimensions of uncertainty. Chapter 1 of *Perspectives on Thinking, Judging, and Decision Making*, 2011.

[18] Michael Gill, Arthur Spirling. Estimating the severity of the WikiLeaks US diplomatic cables disclosure. *Politi- cal Analysis* 23:2 (2015), 299–305.

[19] John Goebel, Dan Teague. How many taxis? *Consortium for Mathematics and Its Applications* 72, 1999.

[20] Jayanta K Ghosh, Mohan Delampady, Tapas Samanta. *An introduction to Bayesian Analysis: Theory and Methods*. Springer, 2006.

[21] Leo A Goodman. Serial number analysis. *Journal of the American Statistical Association* 47:260 (1952), 622–634. [22] Leo A Goodman. Some practical techniques in serial number analysis. *Journal of the American Statistical Asso- ciation* 49:265 (2954), 97–112.

[23] Carlos Gómez Grajalez, Eileen Magnello, Robert Woods, Julian Champkin. Great moments in statistics. *Signifi- cance* 10:6 (2013), 21–28.

[24] Andrew Hodges. Alan Turing: the enigma. W *Alan Turing: The Enigma*. Princeton University Press, 2014.

[25] Michael Höhle, Leonhard Held. Bayesian estimation of the size of a population. Technical Report 499, LMU Munich, Discussion Paper, 2006.

[26] Rob J Hyndman. Computing and graphing highest density regions. *American Statistician* 50:2 (1996), 120–126. [27] Roger W. Johnson. Estimating the size of a population. *Teaching Statistics* 16:2 (1994), 50–52.

[28] Liran Katzir, Edo Liberty, Oren Somekh. Estimating sizes of social networks via biased sampling.

In *Proceedings of the 20th International Conference on World Wide Web*, ss. 597–606, 2011.

[29] Karl-Rudolf Koch. *Introduction to Bayesian Statistics*. Springer, 2007.

[30] John Kruschke. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press, 2014.

[31] Anthony Lee, Steven J. Miller. Generalizing the German tank problem. *PUMP Journal of Undergraduate Research* (2023), 59–95.

[32] Luana Micallef, Pierre Dragicev Jean-Daniel Fekete. Assessing the effect of visualizations on Bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics* 18:12 (2012), 2536–2545.

[33] Frederick Mosteller. *Fifty Challenging Problems in Probability with Solutions*. Courier Corporation, 1987.

[34] John H Mott, Margaret L. McNamara, Darcy M. Bullock. Estimation of aircraft operations at airports using nontraditional statistical approaches. In *2016 IEEE Aero- space Conference*, ss. 1–11. IEEE, 2016.

[35] Kevin P. Murphy. *Probabilistic Machine Learning: An Introduction*. MIT Press, 2022.

[36] James D. Nichols. Capture–recapture models. *BioScience* 42:2 (1992), 94–102.

[37] Alvitta Ottley, Blossom Metevier, P. K. Han, Remco Chang. Visually communicating Bayesian statistics to lay- persons. In *Technical Report*. Tufts University, 2012.

[38] William D. Pearse, Charles C. Davis i in. A statistical estimator for determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution* 1:12 (2017), 1876–1882.

[39] Ilaria Prosdocimi. German tanks and historical records: the estimation of the time coverage of ungauged extreme events. *Stochastic Environmental Research and Risk Assess- ment* 32:3 (2018), 607–622.

[40] Ronald L. Rivest i in. Chaffing and winnowing: Confi-dentiality without encryption. *CryptoBytes (RSA Laboratories)* 4:1 (1998), 12–17.

[41] Harry V. Roberts. Informative stopping rules and inferences about population size. *Journal of the American Statistical Association* 62:319 (1967), 763–775.

[42] David L. Roberts, Andrew R. Solow. When did the dodo become extinct? *Nature* 426:6964 (2003), 245.

[43] W. J. Rosenberg, J. J. Deely. The horse-racing problem, a Bayesian approach. *American Statistician* 30:1 (1976), 26–29.

[44] Richard Ruggles, Henry Brodie. An empirical approach to economic intelligence in World War II. *Journal of the American Statistical Association* 42:237 (1947), 72–91.

[45] Richard L Scheaffer, Ann Watkins, Mrudulla Gnana- desikan, Jeffrey Witmer. *Activity-Based Statistics: Student Guide*. Springer, 2013.

[46] Rens van de Schoot, Sarah Depaoli, Ruth King i in. Bayesian statistics and modelling. *Nature Reviews Methods Primers* 1:1 (2021), 1–26.

[47] Robin Senge, Stefan Bösner, Krzysztof Dembczyński i in. Reliable classification: learning classifiers that distinguish aleatoric and epistemic uncertainty. *Information Sciences* 255 (2014), 16–29.

[48] Aaron Tenenbein. The racing car problem. *American Statistician* 25:1 (1971), 38–40.

[49] Wolfgang Von der Linden, Volker Dose, and Udo Von Toussaint. *Bayesian Probability Theory: Applications in the Physical Sciences*. Cambridge University Press, 2014.

[50] Xiaohan Wu, Margaret E. Roberts, Rachel E. Stern, Benjamin L. Liebman, et al. Augmenting serialized bureaucratic data: the case of Chinese courts. *21st Century China Center Research* 11, 2022.

[51] Haibao Tang, Ewen F Kirkness, et al. Profiling of short-tandem-repeat disease alleles in 12,632 human whole genomes. *American Journal of Human Genetics* 101:5 (2017), 700–715.

Przekład Ernest Aleksy Bartnik
(*Wydział Fizyki UW*)