Statystyka i niepewność: rola danych w nowoczesnym biznesie

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Artykuł stanowi głęboką analizę roli statystyki we współczesnym świecie biznesu, wykraczając poza proste techniki obliczeniowe. Autor, odwołując się do myśli Petera Kenny'ego, przedstawia statystykę jako klucz do zrozumienia antropologii nowoczesnych decyzji podejmowanych w warunkach niepewności. Tekst szczegółowo omawia triadę próbka-populacja oraz metodologię doboru losowego, wskazując na etyczne aspekty dataizmu i odpowiedzialność menedżerską. Czytelnik dowie się, jak unikać typowych pułapek prognozowania, takich jak błędy ekstrapolacji czy nadmierne zaufanie do modeli Gaussa. Całość osadzona jest w kontekście globalnych regulacji dotyczących sztucznej inteligencji i Big Data, promując świadome i odpowiedzialne podejście do analizy danych liczbowych i kategorycznych w procesach zarządczych.

This article provides an in-depth analysis of the role of statistics in the contemporary business world, going beyond simple computational techniques. Drawing on the work of Peter Kenny, the author presents statistics as a key to understanding the anthropology of modern decisions made under conditions of uncertainty. The text thoroughly discusses the sample-population triad and random sampling methodology, highlighting the ethical aspects of dataism and managerial responsibility. The reader will learn how to avoid common forecasting pitfalls, such as extrapolation errors and overreliance on Gaussian models. The entire text is set in the context of global regulations regarding artificial intelligence and big data, promoting an informed and responsible approach to the analysis of numerical and categorical data in management processes.

Artykuł analizuje współczesne gospodarki przez pryzmat niepewności i ryzyka, stawiając tezę, że statystyka, wbrew powszechnemu przekonaniu, nie jest jedynie zbiorem skomplikowanych metod, lecz przede wszystkim sztuką radzenia sobie z niepewnością. Opierając się na myśli Petera Kenny'ego, autor argumentuje, że kluczowe jest zrozumienie ograniczeń danych, procesu próbkowania i ryzyka błędów, aby uniknąć iluzji pewności. Tekst dekonstruuje mit dataizmu,

ukazując, jak nadmierne zaufanie do danych i algorytmów może prowadzić do korporacyjnej schizofrenii i przerzucania odpowiedzialności. Analizie poddane zostają różne podejścia do danych w Ameryce Północnej, Europie i krajach arabskich, a także pułapki związane z prognozowaniem i utożsamianiem korelacji z przyczynowością. Artykuł wzywa do etycznej refleksji nad wykorzystaniem statystyki, podkreślając konieczność kształcenia w duchu statystycznej pokory i wzmacniania instytucji strzegących jakości danych, aby statystyka stała się narzędziem racjonalnej debaty, a nie jedynie językiem władzy.

SŁOWA KLUCZOWE

statystyka | niepewność | dane liczbowe | dane kategoryczne | próbka | populacja | dobór losowy | dataizm | sztuczna inteligencja | modele predykcyjne | błąd losowy | rozkład normalny | ekstrapolacja | Big Data | odpowiedzialność biznesowa

Niepewność: fundament i racja bytu statystyki

Jeśli chcemy zrozumieć współczesne gospodarki nie jako kronikę zdarzeń, lecz jako arenę, gdzie podejmuje się ryzykowne decyzje w zalewie danych, musimy zacząć od zdania Petera Kenny'ego. Jego teza, z pozoru banalna, jest w istocie radykalna: metody statystyczne mogą być skomplikowane, lecz znaczenie statystyki – nie. Ten lapidarny sąd stanowi w rzeczywistości klucz do antropologii nowoczesnych decyzji i struktury naszej samowiedzy. Statystyka, pojmowana poważnie, nie jest bowiem kolekcją technik, lecz sztuką adekwatnego obchodzenia się z niepewnością – stałą cechą świata społecznego, przyrodniczego i gospodarczego. Dopiero pogodzenie się z faktem, że całkowita eliminacja niepewności byłaby równoznaczna z likwidacją wolności i innowacji, pozwala dostrzec w statystyce narzędzie rozumu praktycznego, a nie tylko zaawansowaną formę rachunkowości.

Niepewność nie jest przeciwieństwem racjonalności, lecz jej materiałem. Statystyka przetwarza surowe fakty, zebrane w warunkach nieuchronnej niedoskonałości, w twierdzenia o prawdopodobieństwie. W ten sposób przekształca chaos danych w zdania o tym, jak bardzo możemy polegać na uporządkowanych fragmentach rzeczywistości, gdy koordynujemy nasze działania. Ostateczne wnioski mają zatem postać warunkową: przy takim sposobie pomiaru, przy takiej procedurze losowania i przy takim modelu zależności możemy z określonym marginesem błędu przyjąć, że dana decyzja jest bardziej rozumna niż jej alternatywa. To nie jest obietnica metafizycznej pewności. To jest propozycja rozsądnego hazardu.

Porównanie statystyki do eleganckiej rozrywki, na wzór krzyżówek czy sudoku, nie jest ze strony Kenny'ego kokieterią. W tej metaforze kryje się bowiem ważna intuicja: prawdziwa trudność nie polega na wyrafinowanych rachunkach, lecz na rozpoznaniu, co naprawdę jest niewiadomą w danej sytuacji i które informacje są w ogóle relewantne. To jest właśnie punkt przecięcia między formalną aparaturą a przeżywanym światem menedżera, ekonomisty czy regulatora. Statystyka zaczyna się tam, gdzie ktoś ma odwagę postawić pytanie: czy moje przeczucie ma jakiekolwiek zakotwiczenie w danych, czy jest jedynie artefaktem logiki, która boi się przyznać do własnej niewiedzy?

Dane opisowe vs liczbowe: jakość kontra ilość

Aby wykroczyć poza poziom ogólnych deklaracji, należy powrócić do elementarnej, a często niedocenianej triady: dane, próbka, populacja. Dane, w rozumieniu statystycznym, nie są bowiem czystym zapisem faktów, lecz konstruktem, który już na etapie rejestracji zostaje przefiltrowany przez gęstą siatkę definicji, ograniczeń pomiaru, interesów zbierającego i nieuniknionych kompromisów. Fundamentalne rozróżnienie, którego odtworzenie jest tu niezbędne, przebiega między danymi opisowymi a liczbowymi. Te pierwsze, zwane też kategorycznymi, pozwalają się zorganizować w skończoną liczbę kategorii i policzyć, ile obserwacji przypada na każdą z nich – jak kolor oczu, zawód czy preferencja polityczna sprowadzona do etykiety na liście. Dane liczbowe z kolei mogą być ciągłe, gdy ich wartości układają się na potencjalnie nieprzerwanej skali, jak wzrost czy czas, albo dyskretne, gdy dają się policzyć w jednostkach niepodzielnych, jak liczba dzieci w gospodarstwie domowym.

Statystyka nie pracuje jednak na danych w sensie czysto ontologicznym, lecz na próbkach. Próbka jest skończonym zbiorem obserwacji, pozyskanym w spójnym procesie, który ma ambicję reprezentować jakąś większą całość, zwaną populacją. Populacja zaś, w sensie ścisłym, pozostaje często tworem hipotetycznym i potencjalnie nieskończonym – zbiorowością wszystkich możliwych klientów, wszystkich potencjalnych transakcji czy wszystkich przyszłych odczytów czujnika w silniku odrzutowym. W tej mierze, w jakiej próbujemy z tego, co policzyliśmy w próbce, wnioskować o tym, co zachodzi w populacji, wprowadzamy niepewność z definicji. Niepewność nie jest tu jednak przeciwieństwem racjonalności, lecz jej materiałem. Każda próbka jest bowiem tylko jedną z wielu możliwych, które moglibyśmy uzyskać, gdybyśmy powtórzyli proces losowania nieskończoną liczbę razy.

Metody próbkowania determinują rzetelność wyników

Tu właśnie objawia się pierwszy wymiar statystycznej odpowiedzialności. O tym, czy wyniki nadają się do jakiegokolwiek uzasadnienia, nie decyduje bowiem złożoność rachunków, lecz fundamentalny akt doboru

próby. Prosty dobór losowy, w którym każda jednostka populacji ma jednakową szansę wyboru, pozostaje ideałem, do którego z trudem aspirujemy. Dobór systematyczny, polegający na wyborze co n-tej pozycji z listy, jest wygodnym przybliżeniem, lecz niesie ryzyko, gdy sam porządek listy skrywa własny, niewidoczny rytm. Z kolei dobór warstwowy, świadomie odtwarzający strukturę populacji, oraz zespołowy, losujący całe grupy, to narzędzia nie tyle matematyczne, co instytucjonalne; wymagają one wiedzy o realnej tkance świata społecznego, zanim w ogóle przystąpimy do jego mierzenia.

W praktyce biznesowej ten poziom metodologicznej refleksji bywa jednak lekceważony z uderzającą nonszalancją. Ankiety internetowe namaszcza się na głos ludu, mimo ich wątpliwej reprezentatywności; dane transakcyjne z jednego kanału sprzedaży stają się obrazem całego rynku; obserwacje z kilku sezonów rzutuje się w nieskończoność. W ten sposób niepewność, którą można by uczciwie skwantyfikować, zostaje wygnana na poziom przedrefleksyjnego tła. Wraca jednak nieuchronnie, tyle że już w postaci kryzysu zaufania do samych danych, które miały przecież tę niepewność poskromić.

Paradoks próbki: większa grupa nie gwarantuje prawdy

Relacja między wielkością próby a rzetelnością wniosków pozostaje uporczywą pułapką poznawczą, w którą wpadają nawet profesjonaliści. Intuicja podsuwa nam złudzenie, że próba musi rosnąć wraz z populacją, tymczasem błąd losowy zależy od bezwzględnej liczby obserwacji, a nie od jej proporcji względem całości. Prowadzi to do paradoksalnego, lecz logicznie nieuchronnego wniosku: sondaż tysiąca osób opisuje preferencje miliona wyborców z niemal taką samą precyzją jak preferencje dziesięciu milionów, o ile mechanizm losowania jest rzetelny.

Trzy poziomy niepewności w ujęciu Petera Kenny'ego

Kenny jest w tym punkcie bezlitosny: niepewność nie jest tylko rachunkiem prawdopodobieństwa, jest też rachunkiem sumienia. Błędy pomiarowe, nieprecyzyjne definicje, uprzedzenia badacza, interesy fundatora, niechęć respondentów czy obojętność użytkownika danych na warunki ich powstania – wszystko to tworzy gęstą warstwę niepewności. Warstwa ta nie znika tylko dlatego, że ktoś na końcu wygenerował wykres w wysokiej rozdzielczości. Źródła tej niepewności piętrzą się na co najmniej trzech poziomach.

Po pierwsze, dane pierwotne mogą być po prostu fałszywe lub niedokładne, co dotyczy zarówno ankiet, w których respondenci odpowiadają „tak, jak wypada”, jak i oficjalnych statystyk, które rodzą się w instytucjach mających konkretny interes polityczny lub budżetowy. Po drugie, sam proces przetwarzania danych może wprowadzać dalsze zniekształcenia poprzez nadmierne uśrednianie, zaokrąglanie czy pomijanie obserwacji odstających, które są kłopotliwe dla przyjętej narracji. Wreszcie, operacja przejścia od

próbki do całej populacji wykorzystuje modele, które mówią więcej o naszych założeniach niż o świecie, jeśli tylko nie wypowiemy tych założeń wprost.

W tym kontekście powraca jedna z najbardziej drapieżnych intuicji Florence Nightingale, która, traktując statystykę jako rodzaj świeckiego kultu, pisała, że pozwala ona odczytać ukryty porządek świata społecznego. Dziś trzeba tę intuicję odwrócić w trybie krytycznym. Statystyka pozwala nie tylko odczytywać porządek, lecz także doskonale maskować jego pęknięcia, gdy aparat liczbowy służy do legitymizowania decyzji podjętych wcześniej w obrębie władzy ekonomicznej lub politycznej. Wówczas niepewność nie jest komunikowana jako margines błędu, lecz wypierana, a tym samym obiektywizowana jako rzekoma konieczność.

Trzy tezy o pewności demaskują złudną naturę danych

Spróbujmy przeprowadzić eksperyment myślowy, który na pozór dotyczy jedynie logiki, a w istocie obnaża fundamentalne napięcie w korporacyjnej wierze w dane. Rozważmy trzy zdania wypowiadane z perspektywy zarządu, który właśnie analizuje najnowsze prognozy statystyczne. Po pierwsze: jeżeli nasze dane są w pełni reprezentatywne, to nasze prognozy są nieuchronnie pewne. Po drugie: jeżeli nasze prognozy są pewne, to nie ponosimy odpowiedzialności za wynik, gdyż rzeczywistość musiała się im podporządkować. Po trzecie: jeżeli jednak ponosimy odpowiedzialność za wynik, to nasze dane nie mogły być całkowicie reprezentatywne.

Zadajmy sobie kluczowe pytanie: czy te trzy przekonania mogą być jednocześnie prawdziwe? W kategoriach formalnych wygląda to na prosty test spójności logicznej. Jednak dopiero gdy odczytamy je w kontekście realnych działań biznesowych, ujawniają one głęboką aporię, czyli nierozwiązywalną sprzeczność, nowoczesnego dataizmu. Jeśli przyjmiemy pierwsze zdanie za wyraz naiwnego realizmu, uznając, że idealna próbka gwarantuje pewność, drugie zdanie sprowadza zarząd do roli notariusza konieczności. W takim świecie porażka nigdy nie byłaby skutkiem błędnej decyzji, a jedynie efektem zewnętrznych szoków, których model nie mógł przewidzieć.

Tymczasem trzecie zdanie gwałtownie odwraca perspektywę. Przypomina, że odpowiedzialność pojawia się właśnie tam, gdzie kończy się pewność – ponieważ dane nigdy nie są doskonałym odbiciem świata, decyzja menedżerska zawsze zawiera element aktu woli, którego nie da się zredukować do czystego algorytmu. Rozwiązanie naszej zagadki polega na uświadomieniu sobie, że ten zestaw przekonań jest wewnętrznie sprzeczny. Sprzeczność ta odzwierciedla niemożliwe marzenie o statystyce jako wygodnym substytucie etyki. Nie da się bowiem jednocześnie twierdzić, że dane dają pewność, pewność zwalnia z odpowiedzialności, a odpowiedzialność wynika z niepewności danych.

Kto próbuje tak myśleć, popada w sprzeczność performatywną, czyli sam sobie przeczy w działaniu. W praktyce gospodarczej oznacza to, że menedżer, który z jednej strony powołuje się na nieomylność modeli predykcyjnych, a z drugiej surowo egzekwuje od zespołu odpowiedzialność za wyniki, tworzy kulturę korporacyjnej schizofrenii. W takim środowisku ani analityka, ani etyka nie mogą zdrowo funkcjonować, stając się jedynie narzędziami do przerzucania winy.

USA, UE i kraje arabskie: trzy modele rozwoju AI

Rozdźwięk między obietnicą a realną praktyką staje się szczególnie wyraźny, gdy przyjrzymy mu się przez pryzmat geografii. W Ameryce Północnej, zwłaszcza w Stanach Zjednoczonych, dataizm przyjmuje formę radykalnego pragmatyzmu. Dane traktuje się tam jako zasób strategiczny, często cenniejszy od materialnych aktywów, a gotowość do agresywnego eksperymentowania z modelami AI jest niemal powszechna. To stamtąd płyną raporty, wedle których niemal wszystkie duże firmy inwestują w sztuczną inteligencję, lecz zaledwie ułamek z nich osiąga etap, w którym staje się ona krwiobiegiem codziennych decyzji, a nie tylko efektownym projektem pilotażowym. Jednocześnie rośnie liczba spółek giełdowych, które w oficjalnych dokumentach opisują AI nie tylko jako szansę, lecz jako poważne źródło ryzyka – prawnego, reputacyjnego i operacyjnego, porównywalnego z turbulencjami, jakie niosła ze sobą rewolycja internetowa.

W Unii Europejskiej dominuje odmienna logika. Tutaj historyczna pamięć o nadużyciach władzy nad informacją, połączona z tradycją państwa opiekuńczego i silnych instytucji, zmusza do ujęcia danych w ścisły gorset regulacyjny. Ogólne rozporządzenie o ochronie danych, projektowane akty dotyczące sztucznej inteligencji czy polityki konkurencji wymierzone w cyfrowych gigantów tworzą środowisko, w którym Big Data i AI muszą od początku wpisać się w horyzont etyczny i prawny. Spowalnia to niektóre procesy komercjalizacji, ale jednocześnie zmusza firmy do poważnego traktowania prywatności, przejrzystości modeli i odpowiedzialności za algorytmiczne decyzje. Badania pokazują, że dla wielu europejskich przedsiębiorstw to właśnie niepewność regulacyjna i obawa przed sankcjami stanowią dziś główny hamulec, nawet jeśli towarzyszy im świadomość utraty przewagi konkurencyjnej.

Gospodarki arabskie, szczególnie te, które przeszły gwałtowną dywersyfikację od ropy naftowej ku technologii, działają na jeszcze innym tle kulturowym. Z jednej strony zakorzeniona jest tam głęboka tradycja personalistyczna, w której zaufanie buduje się poprzez sieci pokrewieństwa, patronatu i honorowych zobowiązań. Z drugiej strony elity tych państw intensywnie inwestują w Big Data i AI jako narzędzia modernizacji, nadzoru i zarządzania ryzykiem gospodarczym. Statystyka staje się tam instrumentem hybrydowym, służącym jednocześnie do optymalizacji logistyki i finansów oraz do monitorowania lojalności społecznej. Niepewność jest tu redukowana nie tylko przez modelowanie, lecz także przez wzmacnianie struktur hierarchicznych, które dystrybuują ryzyko polityczne.

Pułapka Malthusa: błędy naiwnej ekstrapolacji trendów

Kenny poświęca istotną część swojej refleksji prognozowaniu, pod którą to nazwą kryje się zarówno rozsądne oczekiwanie ciągłości, jak i niebezpieczne złudzenie, że przyszłość musi być prostą ekstrapolacją przeszłości. Każda matematyczna prognoza opiera się na jednym z dwóch założeń. Albo przyjmujemy, że przyszłe dane będą podlegać temu samemu rozkładowi co dane historyczne, co sprawdza się w przypadku zjawisk rządzonych przez dobrze opisane prawa fizyki. Albo próbujemy dopasować linię prostą lub krzywą do szeregu czasowego, a następnie przedłużamy ją w przyszłość, licząc na to, że świat nie zmieni reguł gry. Obie metody są niezbędne do zarządzania firmą, budżetem państwa czy systemem emerytalnym, lecz obie noszą w sobie zarodek katastrofy, gdy zapominamy o ich warunkach brzegowych.

Przykład Malthusa pozostaje tu niezmiennie pouczający. Jego doktryna, wedle której populacja rośnie geometrycznie, a produkcja żywności arytmetycznie, jawiła się w swoim czasie jako nieuchronna zapowiedź globalnego głodu. Tymczasem rewolucja technologiczna, zmiana wzorców reprodukcyjnych i globalizacja gospodarki unieważniły założenia, które nadawały modelowi Malthusa sens. Ekstrapolacja zawiodła nie dlatego, że rachunek był błędny, lecz dlatego, że funkcja opisująca świat przestała mieć do niego zastosowanie. Równanie było poprawne, ale opisywało rzeczywistość, która przestała istnieć.

Na przeciwległym biegunie leży rozkład normalny, który w praktyce gospodarczej bywa fetyszem, ponieważ oferuje kusząco symetryczny obraz świata, gdzie odchylenia są regularne, a skrajne zdarzenia niemal niemożliwe. Gdy producent odzieży używa danych o wzroście rekrutów, by oszacować zapotrzebowanie na rozmiary, krzywa Gaussa jest narzędziem zdrowego rozsądku. Gdy jednak fundusz inwestycyjny traktuje ten sam rozkład jako opis zmienności rynków finansowych – podatnych na panikę, szoki i nagłe zmiany nastrojów – ta sama matematyka staje się źródłem katastrofalnego niedoszacowania ryzyka. Staje się aktem wiary wbrew faktom.

Współczesny biznes, próbując łączyć prognozowanie statystyczne z narzędziami sztucznej inteligencji, generuje nową jakość i nową skalę problemów. Badania nad AI i Big Data pokazują, że modele uczące się na ogromnych zbiorach danych potrafią radykalnie poprawić trafność prognoz, zwłaszcza w dynamicznym otoczeniu. Jednocześnie te same analizy wskazują na niepokojący paradoks: im modele stają się bardziej złożone i nieprzejrzyste, tym trudniej menedżerom zrozumieć, jakie założenia leżą u ich podstaw. Gdzie kończy się niepewność wynikająca z niedoskonałości modelu, a gdzie zaczyna się ta fundamentalna, nieusuwalna niepewność samego świata?

Wracamy tym samym do intuicji Kenny’ego, który przypomina, że statystyka jest sztuką zadawania właściwych pytań, a nie kultem skomplikowanych odpowiedzi. Sztuczna inteligencja może ograniczyć niepewność wynikającą z ludzkich błędów poznawczych, ale nie zniesie fundamentalnej nieprzewidywalności systemów społecznych, w których decyzje jednych są reakcją na przewidywania

drugich. W tym sensie globalny biznes, który dziś ogłasza AI strategicznym priorytetem, a jednocześnie przyznaje, że większość jego danych jest fragmentaryczna lub niewiarygodna, stanowi żywy komentarz do tezy, iż metody mogą być wyrafinowane, lecz znaczenie jest proste. A znaczenie polega na tym, że bez dobrej próbki, uczciwej definicji zmiennych i pokory wobec niepewności, żadna moc obliczeniowa nie zastąpi rozsądku.

Błędy I i II rodzaju paraliżują decyzje zarządcze

Dopiero gdy wyzbędziemy się iluzji, że dane mówią same za siebie, i przyjmiemy, że zawsze przemawiają w języku, w którym ktoś układa z nich opowieść o świecie, staje się jasne, dlaczego statystyka nie może istnieć bez porównań, hipotez i świadomie przyjętego ryzyka błędu. To właśnie w tym momencie wiedza przestaje być bezdusznym magazynem liczb, a staje się praktyką uzasadniania twierdzeń. Peter Kenny, systematyzując zagadnienie porównań, w istocie pokazuje, jak ucywilizować odwieczną ludzką skłonność do ferowania wyroków na podstawie podobieństw i różnic, które zmysły podsuwają nam w sposób zupełnie bezrefleksyjny.

Porównanie w sensie statystycznym nie jest bowiem prostym zestawieniem dwóch wartości. To precyzyjna procedura, w której najpierw definiujemy, co podlega ocenie, a następnie formułujemy hipotezę, że między porównywanymi wielkościami nie ma żadnej rzeczywistej różnicy. Kluczowe jest jednak to, że z góry ustalamy poziom ryzyka, jaki jesteśmy gotowi ponieść, gdybyśmy tę hipotezę odrzucili. Hipoteza zerowa, zazwyczaj przybierająca formę stwierdzenia, że dwie średnie są równe lub że dwie proporcje nie różnią się istotnie, jest w istocie wyrazem metodologicznej ostrożności. Mówimy przez nią: dopóki dowody na zmianę nie będą przytłaczające, za bardziej racjonalne uznajemy założenie o braku zmiany.

Poziom istotności, na przykład owe słynne pięć procent, staje się w tym ujęciu kwantyfikacją intelektualnej uczciwości. Przyjmując taki próg, zobowiązujemy się, że tylko w jednym przypadku na dwadzieścia pozwolimy sobie ogłosić istnienie różnicy, jeśli dane pochodzące ze świata, w którym nic się nie zmieniło, przypadkiem ułożą się w obraz pozornej zmiany. Błąd pierwszego rodzaju, czyli fałszywy alarm, nie jest tu wypadkiem przy pracy, lecz wkalkulowanym ryzykiem. Statystyka nie eliminuje pomyłki, lecz sprawia, że przestaje ona być owocem kaprysu, a staje się elementem kontrolowanego rachunku.

W tym miejscu jednak rodzi się fundamentalne napięcie, które Kenny traktuje z całą powagą, a które praktyka biznesowa często ignoruje. Jeśli bowiem obniżymy próg istotności, żądając na przykład jednego procenta, aby chronić się przed pochopnym ogłaszaniem zmian, to automatycznie zwiększamy prawdopodobieństwo błędu drugiego rodzaju. Ten błąd polega na przegapieniu różnicy, która faktycznie istnieje. W realiach gospodarczych oznacza to, że możemy zignorować krytyczne symptomy kryzysu,

przeoczyć zmianę nastrojów klientów czy zlekceważyć wzrost awaryjności produktu. Innymi słowy, ceną za mniejszą liczbę fałszywych alarmów jest większa liczba przeoczonych zagrożeń.

Istotność statystyczna a realny zysk ekonomiczny

Wyobraźmy sobie dyrektora finansowego średniej firmy produkcyjnej, który otwiera raport analityka. Informacja jest z pozoru jednoznaczna: średni zysk z nowej linii produktów jest o trzy procent wyższy od dotychczasowego. Bez statystyki, a raczej bez zrozumienia jej ducha, reakcja jest odruchowa i, co gorsza, banalna. Trzy procent więcej to trzy procent więcej. Należy zatem natychmiast przerzucić zasoby na nową linię, a starą metodycznie wygasić. W praktyce tak właśnie często się dzieje.

Analityk, który rozumie istotę rzeczy, zrobi jednak coś innego. Zada pytanie o rolę losowości i o to, na ile obserwowane trzy procent mieszczą się w naturalnym rozrzucie wyników. Sięgnie po błąd standardowy średniej – miarę, która konfrontuje wariancję danych z liczbą obserwacji – by odpowiedzieć, czy różnica między liniami jest większa niż to, czego moglibyśmy się spodziewać, gdyby w istocie obie były równie zyskowne. Jeśli okaże się, że różnica trzech procent jest mniejsza niż na przykład dwukrotność błędu standardowego, rozsądny dyrektor powstrzyma się przed radykalnymi ruchami. Statystyka, daleka od paraliżu decyzyjnego, staje się tu narzędziem obrony przed własnym, kosztownym entuzjazmem.

W ten sposób statystyka nie zastępuje osądu, lecz wzywa go do samoograniczenia. Jeżeli wynik jest istotny statystycznie, nie oznacza to automatycznie, że jest istotny ekonomicznie. Wystarczająco duża próba sprawi, że nawet minimalne różnice w średnich wykażą statystyczną istotność. Wówczas rolą decydenta jest rozstrzygnięcie, czy zysk rzędu dwóch dziesiątych procenta ma jakikolwiek sens praktyczny w obliczu ogromnych kosztów reorganizacji. W stopniu, w jakim świat biznesu zapomina o tym rozróżnieniu, wpycha statystykę w rolę kozła ofiarnego. Gdy decyzja okaże się nietrafiona, winna jest rzekomo liczba, a nie sposób, w jaki została zinterpretowana.

Humor tej abstrakcji objawia się w anegdotycznej figurze prezesa, który domaga się od działu analiz prognozy z dokładnością do dwóch miejsc po przecinku, po czym podejmuje decyzję na podstawie chwilowego przeczucia. To nie jest tylko dowcip. To jest portret logicznej sprzeczności, w której paradygmat samotnej, przekonanej o własnej nieomylności refleksji posługuje się językiem statystyki wyłącznie jako dekoracją, a nie jako medium samokontroli.

Korelacja to nie przyczynowość: błędy analiz rynkowych

Przechodząc od prostych porównań do związków między zmiennymi, wkraczamy w obszar, w którym nasza codzienna intuicja najczęściej prowadzi nas na manowce. Umysł ludzki jest bowiem maszyną do poszukiwania przyczyn, toteż widząc dwie rzeczy zmieniające się w podobnym rytmie, niemal odruchowo zakłada, że jedna musi wynikać z drugiej. Kenny, podążając za klasyczną mądrością metodologiczną, powtarza powściągliwe, lecz fundamentalne zdanie: korelacja jest warunkiem koniecznym przyczynowości, ale nigdy wystarczającym. Bez związku nie ma mowy o przyczynie, lecz sam związek niczego jeszcze nie dowodzi.

Współczynnik korelacji Pearsona, przyjmujący wartości od -1 do +1, jest jedynie termometrem mierzącym temperaturę liniowego związku między dwiema zmiennymi liczbowymi. Gdy zbliża się do jedynki, punkty na wykresie układają się w niemal idealnie prostą, rosnącą linię. Gdy oscyluje wokół zera, nie widać żadnego wyraźnego trendu. Cała sztuka polega na tym, by nie wyczytywać z tej liczby więcej, niż ona faktycznie komunikuje. Silna dodatnia korelacja między liczbą sprzedanych parasoli a liczbą zamówień na pizzę nie oznacza przecież, że schronienie się przed deszczem wywołuje apetyt. Obie zmienne są raczej skutkiem trzeciego, ukrytego czynnika – niepogody.

Modelowym przykładem takiej pozornej zależności, który przywołuje Kenny, jest słynna historia o bocianach i narodzinach dzieci w Kopenhadze. Wzrost liczby gniazdujących ptaków zdawał się iść w parze ze wzrostem liczby urodzeń, co stało się pożywką dla medialnych nagłówków, próbujących ocalić starą baśń w nowym, statystycznym języku. Tymczasem znacznie bardziej prozaiczne wyjaśnienie odwołuje się do trzeciej zmiennej: powojennego boomu budowlanego. Dostarczył on jednocześnie więcej mieszkań dla ludzi i więcej kominów dla bocianów. Korelacja nie jest więc śladem magicznego połączenia, lecz sygnałem, że coś łączy zmienne. Zadaniem analityka jest dociekać, czym jest to „coś”.

Współczesna myśl ekonomiczna, od Judei Pearla po badaczy polityki publicznej, rozwija teorię przyczynowości opartą na modelach strukturalnych i kontrfaktycznych, które próbują zbudować most między korelacją a ugruntowanym twierdzeniem o przyczynie. Idea jest prosta w swej ambicji, lecz niezwykle trudna w realizacji. Chodzi o skonstruowanie takiego modelu zależności, który pozwoliłby odpowiedzieć na pytanie: co by się stało, gdybyśmy zmienili jedną zmienną, utrzymując pozostałe bez zmian? Statystyka klasyczna, którą opisuje Kenny, dostarcza tu niezbędnego punktu wyjścia, ale sama nie wystarcza, by udzielić odpowiedzi na tak postawione pytanie. Nie jest przestarzała, lecz musi zostać wprzęgnięta w aparat bardziej wymagających teorii.

W świecie biznesu pokusa utożsamienia korelacji z przyczynowością jest szczególnie niebezpieczna w marketingu i polityce cenowej. Menedżerowie, widząc wzrost sprzedaży w okresie intensywnej kampanii reklamowej, są skłonni uznać ją za sukces. Mogli jednak przeoczyć równoczesną zmianę cen, sprzyjającą pogodę lub ważne wydarzenie społeczne. Tu właśnie ujawnia się sens badań kontrolowanych i testów A/B. Jeśli losowo podzielimy klientów na dwie grupy, z których jedna widzi reklamę, a druga nie, to ewentualne różnice w sprzedaży można z o wiele większą pewnością przypisać właśnie kampanii.

Problem w tym, że logika eksperymentu jest trudna do zastosowania w wielu sytuacjach biznesowych, a w sferze polityki publicznej napotyka bariery etyczne. Nie można przecież losowo przydzielać dzieci do szkół o róż

Big Data i AI rewolucjonizują warsztat statystyka

Dopiero gdy wyczerpiemy klasyczny repertuar pojęć – niepewności, danych, próbkowania, porównań, związków i prognoz – możemy spróbować pojąć, co naprawdę kryje się za twierdzeniem, że Big Data i sztuczna inteligencja są najważniejszym rozwinięciem statystyki ostatnich dekad. Peter Kenny słusznie podkreśla, że nie istnieje jedna, precyzyjna definicja Big Data, ponieważ nie chodzi tu o arbitralny próg ilościowy. To raczej sytuacja strukturalna, w której zbiory danych przekraczają zdolności tradycyjnych technologii, a jednocześnie wnikają w sam rdzeń procesów decyzyjnych. Zasadniczą cechą Big Data jest jej charakter nadmiaru w stosunku do klasycznych procedur; nie chodzi już o ostrożne dobieranie prób, ale o oswajanie potopu informacji.

Tradycyjna triada objętości, szybkości i zmienności, którą zwykło się opisywać Big Data, nabiera pełnego sensu dopiero po dodaniu czwartego, kluczowego wymiaru: wiarygodności. Można bowiem zbudować hurtownię danych o pojemności petabajtów, korzystać z chmury obliczeniowej o przepustowości godnej programu kosmicznego i zaprząc do pracy całe klastry serwerów, a mimo to nie uzyskać ani jednego wiarygodnego wniosku. Wystarczy, że dane wejściowe będą zanieczyszczone błędami pomiarowymi, systematycznymi uprzedzeniami czy brakiem precyzyjnych definicji. Kenny, opisując tradycyjną statystykę, uczy pokory wobec każdego z tych zagrożeń. Big Data nie znosi tych wymagań – jedynie mnoży ich skalę.

Dla praktyki gospodarczej szczególne znaczenie ma przejście od analizy opisowej danych do eksploracji predykcyjnej. Eksploracja danych w ścisłym tego słowa znaczeniu nie jest już tylko zestawieniem wskaźników, lecz systematycznym wydobywaniem wzorców o wartości prognostycznej. Analityka predykcyjna, rozumiana jako procedura statystyczna, w której wyniki wyrażane są w kategoriach prawdopodobieństwa, przekształca się w sercu przedsiębiorstw w aparat sterujący logistyką, zarządzaniem ryzykiem, ofertami dla klientów czy polityką kredytową. Modele oparte na drzewach decyzyjnych, klastrowaniu, sieciach neuronowych i zespołach modeli działają już nie w laboratoriach, ale w czasie

rzeczywistym – w systemach scoringowych, silnikach rekomendacji i na algorytmicznych platformach transakcyjnych.

Z perspektywy paradygmatu naukowego zjawisko to można opisać jako przesunięcie akcentu z modeli wywodzących się z teorii na modele wywodzące się z danych. W ekonomii, która przez dziesięciolecia karmiła się równaniami równowagi ogólnej i teoriami racjonalnych oczekiwań, coraz większą rolę odgrywają podejścia empiryczne, oparte na eksperymentach naturalnych i masowym wykorzystaniu danych mikroekonomicznych. W tym pejzażu Big Data i sztuczna inteligencja jawią się jako skrajna forma tego zwrotu ku empirii: zamiast stawiać hipotezy teoretyczne i testować je na ograniczonych zbiorach, pozwalamy algorytmom samodzielnie wyławiać regularności z oceanu obserwacji.

Problem polega na tym, że takie podejście zawiera w sobie poważną aporię. Im bardziej ufamy, że algorytmy znajdą dla nas użyteczne związki, tym mniej jesteśmy skłonni pytać o założenia, które w te algorytmy wbudowano. W rezultacie władza nad definicją niepewności, nad określaniem, co jest istotnym sygnałem, a co jedynie szumem, przesuwa się z poziomu deliberacji między ekspertami w głąb kodu. Niepewność staje się parametrem technicznym, a nie kategorią dyskursu publicznego.

Sztuczna inteligencja i statystyka: nowa dialektyka

Z punktu widzenia Kenny'

Biznes wobec automatyzacji: od entuzjazmu po opór

Wobec dynamicznego rozwoju sztucznej inteligencji w środowiskach gospodarczych wyłaniają się trzy dominujące narracje na temat niepewności, danych i prognozowania. Pierwszą z nich można nazwać technokratycznym optymizmem. W tej wizji AI, karmiona nieustannie rosnącymi zbiorami danych, ma stopniowo doprowadzić do radykalnej redukcji niepewności, czyniąc świat przewidywalnym. Główne zadanie menedżera przyszłości nie będzie już polegało na formułowaniu własnych hipotez, lecz na interpretacji wyników modeli i wdrażaniu ich rekomendacji. Statystyka zostaje tu zredukowana do roli bezszelestnej, niewidzialnej infrastruktury, która po prostu ma działać w tle.

Druga opowieść, znacznie bardziej sceptyczna, podkreśla, że AI co najwyżej przemebluje krajobraz niepewności, lecz go nie usunie. Przedstawiciele tego nurtu, często związani z ekonomią behawioralną i teorią złożoności, argumentują, że im gęściej systemy społeczne oplata sieć reaktywnych podmiotów, tym bardziej prognozowanie przypomina grę luster. To partia szachów, w której każdy z graczy próbuje przewidzieć ruchy pozostałych. W takim świecie AI staje się kolejnym uczestnikiem gry, a nie wszechwiedzącym obserwatorem z zewnątrz, przez co jej obecność wprowadza nowe, nieznane dotąd źródło

nieprzewidywalności. Statystyka przyszłości musi więc uznać, że niepewność nie jest przeciwieństwem racjonalności, lecz jej materiałem.

Trzecia narracja, najbardziej radykalna i krytyczna, zwraca uwagę na fakt, że mariaż AI, Big Data i statystyki staje się narzędziem władzy nad dystrybucją ryzyka. To modele kredytowe decydują, kto otrzyma finansowanie, a kto zostanie z niego wykluczony; algorytmy ubezpieczeniowe ustalają składki w sposób mogący utrwalać historyczne nierówności; systemy scoringu pracowniczego wpływają na to, kto dostanie awans, a kogo uzna się za zbyt ryzykownego. W tym ujęciu statystyka nie jest więc niewinnym narzędziem opisu, ale potężnym instrumentem, który bez świadomego nadzoru normatywnego może reprodukować i pogłębiać strukturalne niesprawiedliwości.

Globalne tendencje, które biznes zdaje się dziś faworyzować, zmierzają w kierunku rosnącej automatyzacji rutynowych decyzji i intensywnego wykorzystania predykcji w zarządzaniu łańcuchami dostaw czy mikrosegmentacji klientów. Analityka predykcyjna wkracza też na obszary dotąd zarezerwowane dla ludzkiego osądu, takie jak rekrutacja, ocena ryzyka prawnego czy diagnoza medyczna. Jednocześnie narasta świadomość, że bez solidnych ram prawnych i etycznych pogoń za predykcyjną doskonałością może wygenerować skutki uboczne, których społeczny koszt dalece przewyższy biznesowe korzyści.

Statystyka jako fundament etyki i odpowiedzialności

Nie ma sposobu, aby ten konflikt rozwiązać czysto technicznie. Możemy wprawdzie zmniejszać oba rodzaje błędów jednocześnie poprzez zwiększanie liczby obserwacji, lecz nawet nieskończona próba nie zlikwiduje fundamentalnej konieczności wyboru progu istotności. I właśnie w tym punkcie statystyka ujawnia swoją etyczną naturę. Decyzja o tym, jaki poziom błędu pierwszego rodzaju jesteśmy gotowi zaakceptować, jest bowiem decyzją o tym, jak rozkładamy odpowiedzialność między tych, którzy reagują zbyt wcześnie, a tych, którzy reagują zbyt późno. W medycynie klinicznej uprzywilejowuje się zwykle ochronę pacjenta przed fałszywie dodatnim wynikiem, stawiając wysokie wymagania dowodowe, zanim dopuści się nowe leczenie. W bezpieczeństwie sejsmicznym czy klimatycznym jest odwrotnie – rozsądniejsze bywa zaakceptowanie większego ryzyka fałszywego alarmu niż pogodzenie się z przegapieniem katastrofy.

Kiedy tego rodzaju logikę stosujemy w zarządzaniu, ujawnia się charakterystyczna hipokryzja kultury korporacyjnej. Zarządy lubią powoływać się na twarde liczby, aby legitymizować brak działania, wymawiając się brakiem istotności statystycznej, gdy chodzi o podwyżki, inwestycje w bezpieczeństwo pracy czy programy środowiskowe. Natomiast przy ogłaszaniu sukcesów marketingowych są skłonne zaaprobować najwątlejsze nawet różnice jako znaczące. Kenny proponuje tu trzeźwe podejście, w którym poziom istotności nie jest zaklęciem rzucanym dla wygody, lecz parametrem świadomie dobranym do

kontekstu i wagi decyzji. Niepewność nie jest tu tylko rachunkiem prawdopodobieństwa, jest też rachunkiem sumienia.

Klasyczne narzędzia, takie jak test Z do porównywania średnich przy znanej zmienności, test t Studenta przy małych próbach, test F do porównywania wariancji czy analiza wariancji ANOVA, która pozwala badać wiele grup jednocześnie, nie są jedynie sztuczkami matematycznymi. Stanowią one formalizację pewnego etosu. Etosu, w którym nie wystarczy powiedzieć, że jedna kampania reklamowa była lepsza od drugiej, bo sprzedaż wzrosła. Należy jeszcze zapytać, jak duża jest niepewność tego osądu, wynikająca z przypadkowych fluktuacji popytu, sezonowości czy zmian w otoczeniu, których nie kontrolowaliśmy.

Kenny i struktury władzy: granice teoretycznego ujęcia

Choć pozostajemy wierni intuicjom i tezom Petera Kenny'ego, nie możemy uchylić się od polemicznej refleksji nad granicami jego ujęcia. Jego perspektywa, głęboko zakorzeniona w doświadczeniu przemysłu i edukacji, koncentruje się na tym, jak statystyka może pomóc jednostce i organizacji w podejmowaniu lepszych decyzji w świecie fundamentalnej niepewności. Można jednak zarzucić tej wizji, że zbyt mało uwagi poświęca architekturze władzy, w której statystyka jest zarówno wytwarzana, jak i wykorzystywana. Niepewność bowiem nie jest demokratyczna i nie rozkłada się symetrycznie. Jedni dysponują zasobami pozwalającymi przekształcić ją w kalkulowane ryzyko, podczas gdy inni są skazani na dryfowanie w sferze czystej niepewności, pozbawieni dostępu do danych i aparatury ich interpretacji.

Można też wskazać, że Kenny, koncentrując się na próbkach, testach i prognozach, bada raczej pionowy wymiar relacji między danymi a decyzją, zaniedbując poziomy wymiar dyskursu publicznego. To właśnie tam wyniki statystyczne są przedstawiane, negocjowane i kwestionowane. Statystyka funkcjonuje w społeczeństwie nie tylko jako narzędzie decyzyjne, ale także jako język sporów, a nierzadko amunicja w walce o definicję rzeczywistości. Dane o nierównościach, ubóstwie, zmianach klimatycznych czy efektywności polityk publicznych stają się częścią tej batalii. W tym sensie trzeba uzupełnić perspektywę Kenny'ego o refleksję nad tym, jak statystyka staje się elementem koordynacji zbiorowych działań, a nie tylko narzędziem jednostkowego osądu.

Te zastrzeżenia nie podważają jednak rdzenia jego stanowiska, lecz raczej wskazują obiecujące kierunki jego rozwinięcia. Jeżeli chcemy, aby statystyka była nie tylko techniczną dyscypliną, ale także elementem kultury krytycznej, musimy wzmocnić jej wymiar refleksyjny. Oznacza to, że obok nauczania metod testowania, regresji czy próbkowania trzeba uczyć także rozumienia źródeł niepewności, struktury błędów, ograniczeń modeli, a nade wszystko etycznych konsekwencji decyzji, które na podstawie statystyki podejmujemy. Niepewność nie jest tylko rachunkiem prawdopodobieństwa, jest też rachunkiem sumienia.

Rozwiązania, które miałyby być zarazem racjonalne i profetyczne, wymagają kilku zintegrowanych ruchów. Po pierwsze, trzeba wzmocnić instytucje strzegące jakości danych: urzędy statystyczne, jednostki audytu, wewnętrzne komórki odpowiedzialne za zarządzanie informacją. Po drugie, należy rozwijać ramy prawne i etyczne dla sztucznej inteligencji, które będą nie tyle katalogiem zakazów, ile proceduralnym szkieletem, w którym modele muszą być wyjaśnialne, audytowalne i otwarte na krytykę. Po trzecie, konieczne jest kształcenie menedżerów, prawników i decydentów w duchu statystycznej pokory, aby przestali traktować liczby jako alibi, a zaczęli jako narzędzie samokontroli.

Wreszcie, po czwarte, trzeba zadbać o to, aby obywatele mieli dostęp do narzędzi pozwalających im rozumieć statystykę na poziomie wystarczającym do krytycznej oceny narracji, które ich dotyczą. W przeciwnym razie statystyka nieuchronnie stanie się hermetycznym językiem kapłanów władzy, a nie wspólnym medium racjonalnej debaty. Kenny, pisząc, że choć metody mogą być skomplikowane, znaczenie statystyki jest proste, w istocie apeluje, byśmy na nowo nauczyli się traktować liczby jako element dialogu. Dialogu o tym, jak żyć w świecie, w którym niepewność jest nieusuwalna, ale nie musi być ślepa. Jeżeli globalny biznes, nauka i polityka przyjmą tę lekcję, statystyka

> Statystyka, pozbawiona etycznego kompasu, staje się jedynie narzędziem do legitymizacji z góry przyjętych decyzji, maskując pęknięcia w iluzji porządku. Czy w świecie, gdzie algorytmy coraz śmielej kreują naszą rzeczywistość, potrafimy odróżnić prawdę od precyzyjnie wyreżyserowanego kłamstwa? Może prawdziwa mądrość statystyki tkwi nie w poszukiwaniu pewności, lecz w umiejętności dostrzegania niepewności, która definiuje nasze człowieczeństwo.

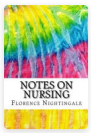
Inspiracje

[1] **"Better business. Decisions from data"** Thomas Robert Malthus

Literatura uzupełniająca

Książki

Literatura pokrewna (Google Scholar):

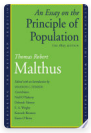


[1] **"Notes on Nursing: What it is and What it is Not"**

Florence Nightingale

2017 | Createspace Independent Publishing Platform | ISBN: 9781548396183

Google Scholar



[2] **"An Essay on the Principle of Population"**

Thomas Robert Malthus

2018 | Yale University Press | ISBN: 9780300231892

Google Scholar



[3] **"Principles of Political Economy"**

Thomas Robert Malthus

2015 | John Murray | ISBN: 9781294954637

Google Scholar



[4] **"Disquisitiones Arithmeticae"**

Carl Friedrich Gauss

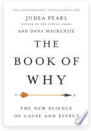
2018 | Springer | ISBN: 9781493975600



[5] **"Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium"**

Carl Friedrich Gauss

2022 | Legare Street Press | ISBN: 9781018079196



[6] **"The Book of Why: The New Science of Cause and Effect"**

Judea Pearl, Dana Mackenzie

2018 | Basic Books | ISBN: 9780465097616

Google Scholar

[7] **"Mortality of the British Army, at home, at home and abroad, and during the Russian war, as compared with the mortality of the civil population in England"**

Florence Nightingale

1858 | Harrison and Sons

Google Scholar



[8] **"Business in Uncertainty: Evolution, Transformation and Adaptation Across Disciplines and Typologies"**

Demetris Vrontis

2025 | Springer Nature | ISBN: 9783031897986

Google Scholar

Artykuły naukowe

Autorzy przywoływani:

[1] "Business Uncertainty in Developing and Emerging Economies"

José M. Barrero, Emilio Zarazúa

2023 | The World Bank

Google Scholar

[2] "Managing uncertainty during a global pandemic: An international business perspective"

Shahzad (Ed) Ansari, Sajjad Yaghoubi, Catherine Welch, Laszlo Tihanyi

2020 | Journal of International Business Studies

Google Scholar

[3] "Sources of Uncertainty in Supervised Machine Learning -- A Statisticians' View"

Cornelia Gruber et al.

2023

Google Scholar

[4] "Theoria combinationis observationum erroribus minimis obnoxiae"

Carl Friedrich Gauss

1823

[5] "General Investigations of Curved Surfaces"

Carl Friedrich Gauss

1827

[6] "Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity and Panmixia"

Karl Pearson

1896 | Philosophical Transactions of the Royal Society of London

[7] "On the Theory of Contingency and its Relation to Association and Normal Correlation"

Karl Pearson

1904

[8] "An introduction to causal inference"

Judea Pearl

2010 | International Journal of Biostatistics

Google Scholar

[9] **"Causal diagrams for empirical research"**

Judea Pearl

1995 | Biometrika

Google Scholar

APA ONE Publishing System

Fundacja Dobre Państwo © 2026

Article ID: 64f288e7-3cdd-4b94-877f-57d228399d0a