

Rozdział 1

Podstawowe pojęcia rachunku prawdopodobieństwa

Rachunek prawdopodobieństwa jest tym działem matematyki, który bada i opisuje zjawiska, uwzględniając ich losowy charakter. Analiza zachowanych do naszych czasów materiałów źródłowych wskazuje, że problematyka o charakterze probabilistycznym podejmowana była w różnorodnych dziedzinach działalności człowieka już wiele lat przed naszą erą. Znane są starożytne dokumenty dotyczące określania liczebności wojska, które zawierają informacje o prognozowaniu liczby chłopców i dziewcząt w populacji, na podstawie danych czerpanych z przeprowadzanych spisów ludności. Pierwszy regularny spis ludności przeprowadzono w Kanadzie (w prowincji Quebec) w 1666. Z kolei w Europie pionierami w tej dziedzinie byli Normanowie i Wilhelm Zdobywca, który w 1066 roku zarządził spis nowo pozyskanych ziem i dóbr. Takie działania spotykały się wówczas z negatywną reakcją. Uważano je wręcz za świętokradztwo, ze względu na opowieść biblijną o królu Dawidzie, który nie mógł przeprowadzić spisu ludności, ponieważ został on przerwany przez zarazę. Regularny spis ludności w Europie został przeprowadzony w Islandii w 1703 roku i w Szwecji w 1749 roku. Innym przykładem zastosowania analizy danych mogą być pochodzące z IX wieku z Indii opracowania zaleceń dla kupców, dotyczące rozmieszczania w różnych agencjach różnych wkładów tak, aby w przyszłości mogli oni czerpać z nich jak największe zyski. Entuzjazm i zapał do regularnego zbierania i opracowywania danych liczbowych dotyczących populacji i gospodarek rozpoczął się dopiero pod koniec XVII wieku, a prawdziwie doceniony został dopiero w pierwszej połowie XIX wieku. Zaczęto porównywać badania społeczne z badaniami np. z zakresu fizyki (na przykład James Maxwell objaśniał swoje prace związane z teorią gazów posługując się statystyką społeczną). Planowanie przestało się opierać na domysłach, a zaczęło w większym stopniu wykorzystywać analizy i obliczenia. Pojawiły się zagadnienia związane z niezależnością i współzależnością zdarzeń i czynnikami mającymi wpływ na otrzymane rezultaty. Okazało się, że zbieranie danych i losowy dobór jednostek do badań jest również złożonym zagadnieniem. Na przykład jednym z sukcesów była technika losowania warstwowego. Zastosował ją w 1936 roku George Gallup do określenia wyników wyborów prezydenckich w USA. Okazało się, że liczna próba nie gwarantuje

jakości badań – próbka musi być reprezentatywna (Literary Digest, konkurent Galupa, posłużył się wynikami badania próby 10 milionów osób, otrzymując fałszywe wyniki, podczas gdy Gallup przebadał zaledwie 3 tysiące, słusznie przewidując na tej podstawie zwycięstwo Franklina Delano Roosevelta w wyborach prezydenckich).

Nie da się ukryć, że do znacznego rozwoju rachunku prawdopodobieństwa przyczynił się początkowo hazard. Pierwsze znane zagadnienia z rachunku prawdopodobieństwa dotyczyły właśnie gier hazardowych, a w szczególności gry w kości, która znana była już w starożytnym Egipcie i Grecji. Wiara, że wygrywa gracz bardziej miły bogom, łączyła się jednocześnie z przekonaniem, że można dopomóc losowi przesuując środek ciężkości kostki. Próby obliczenia możliwych wyników w rzucie kilkoma kostkami można znaleźć w pracach z X i XI wieku. Pierwsze teoretyczne zainteresowanie tą grą wyrazili jednak dopiero matematycy w XVI i XVII wieku. Byli to między innymi Galileusz, Leibniz, Fermat i Pascal. Bohaterem znanej anegdoty jest matematyk Nicolaus Bernoulli, którego sprowadził do Petersburga car Piotr I. Bernoulli zarabiał na balach i przyjęciach przyjmując zakłady – twierdził, że na sali balowej znajdują się dwie osoby, które urodziły się tego samego dnia i miesiąca (dowolnego roku). Jeżeli na sali znajdowało się 31 osób, prawdopodobieństwo przekraczało 50 procent, a każda kolejna obecna osoba oznaczała większą szansę na wygraną.

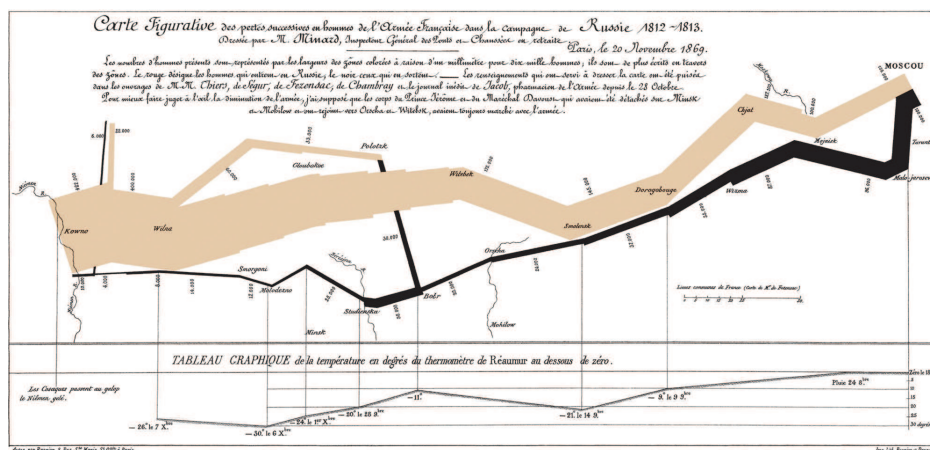
Prawdopodobieństwo stało się precyzyjnym pojęciem dopiero w XVIII wieku. Niestety do XIX wieku traktowano je bardzo ostrożnie. Nawet francuski matematyk de Laplace nazywał je „zdrowym rozsądkiem zredukowanym do obliczeń”.

Historia rachunku prawdopodobieństwa to dzieje przejścia od czasów wyroczni czy wróżbitów do konstrukcji potężnych narzędzi zarządzania ryzykiem i planowania, jakimi dysponujemy dzisiaj. Pierwszy traktat na temat obliczania ryzyka został napisany przez Jana de Wit w 1671 roku w Holandii. Wówczas państwo, aby zebrać pieniądze na przykład na finansowanie wojen, sprzedawało obywatelom renty. Wpłacany corocznie zysk właściciela był niezależny od jego wieku i stanu zdrowia i wynosił jedną siódmą wartości lokaty. De Wit wskazywał na zagrożenie w postaci ewentualnych strat finansowych ponoszonych przez państwo z powodu takiego sposobu wypłacania rent. Dopiero w 1762 roku jedna z londyńskich firm (Equitable) zaczęła brać pod uwagę uprzednio skalkulowane ryzyko przy wycenianiu wartości swoich polis. Rozwój gospodarki kapitalistycznej oraz wzrost liczby transakcji handlowych opartych na kredytowaniu doprowadziły do powstania firm ubezpieczających ryzyko kredytowe. Niestety stosowanie złych zasad przy konstrukcji i zawieraniu ubezpieczeń kredytów było główną przyczyną bankructw wielu firm ubezpieczeniowych powstałych w XIX wieku. Obecnie potrafimy na przykład określić liczbę dzieci przyjętych do szkoły w najbliższym roku szkolnym. Liczba ta z pewnością nie jest równa liczbie dzieci urodzonych sześć lat wcześniej, ani liczbie dzieci kończących przedszkole. Trzeba bowiem uwzględnić dodatkowe czynniki, takie jak śmiertelność w różnych grupach wiekowych, choroby, mechanizmy migracyjne spowodowane na przykład sytuacją ekonomiczną, których zasięg można określić na podstawie różnorodnych obserwacji.

Zdecydowana większość procesów fizycznych, technicznych, biologicznych, społecznych czy ekonomicznych przebiega w sposób bardziej lub mniej losowy. Gdy w prak-

tyce niedostatecznie znamy wszystkie przyczyny decydujące o przebiegu pewnych procesów, to decydujemy się na badanie jak największej liczby przypadków. Staramy się stworzyć model danego zjawiska, będący przybliżeniem rzeczywistości, a tym samym miarą naszej wiedzy na temat planowania i przewidywania przyszłości.

Poniżej zamieszczony jest przykład reprezentacji graficznej danych statystycznych.



Rys. Zagłada Wielkiej Armii Napoleona w Rosji w 1812 r. na wykresie C. J. Minarda – <http://en.wikipedia.org/wiki/File:Minard.png> (Opracowanie: M. Minard, Inspektor Generalny Mostów i Dróg, na emeryturze. Opublikowano w Paryżu, 20 listopada 1869 r.)

Symboliczna mapa, autorstwa francuskiego inżyniera Charlesa Josepha Minarda z 1869 roku, pozwala prześledzić systematyczną zagładę Wielkiej Armii Napoleona I w Rosji w 1812 roku. Z kampanii wróciło zaledwie 4% uczestników. Przedstawiony graficznie na planie stan liczebny wojsk skorelowano z temperaturami panującymi podczas wyprawy Napoleona na Rosję. W opisie karty możemy przeczytać, że każdy milimetr szerokości wykresu odpowiada około 10 tys. żołnierzy. Dla ułatwienia periodycznie opisano stan liczebny na marginesie. Różnymi kolorami oznaczono wojska w drodze na Moskwę i odwrót Wielkiej Armii. Informacje o stanie liczebnym zaczerpnięto z prac M. M. Thiersa, Segura, Fezensacka, Chambraya i nieopublikowanego dziennika Jacoba, farmaceuty armii. Mapę opublikowano 20 listopada 1869 roku w Paryżu.

Nawet jeśli nie zwracamy na to uwagi, to w codziennym życiu zachowujemy się w sposób, który można opisać za pomocą modelu statystycznego. Na przykład, gdy płacimy kartą lub wybieramy pieniądze z bankomatu, informacja na temat naszych działań (ile, gdzie, o której godzinie) trafia do systemu komputerowego. Z tych danych tworzony jest statystyczny model klienta. Na jego podstawie można ustalić, ile średnio (dziennie / miesięcznie) wydajemy, w jakich miejscach (sklepach / miastach / krajach) bywamy. Dzięki temu system bankowy reaguje, gdy dzieje się coś nietypowego – pracownik banku kontaktuje się z klientem, gdy na przykład niespodziewanie robi

on duże zakupy w innym kraju (jak Peru) lub nagle tankuje tonę paliwa na drugim końcu Polski. W bankach miesięcznie włącza się kilkadziesiąt tysięcy takich alarmów. Reguły statystyczne służą również do zwalczania spamu (niechcianych e-maili). Specjalnie opracowane oprogramowanie uczy się najczęściej wykorzystywanych w tego typu wiadomościach zwrotów, co pozwala na wyeliminowanie ich ze skrzynki pocztowej.

Podsumujmy – rachunek prawdopodobieństwa zajmuje się badaniem abstrakcyjnych pojęć matematycznych stworzonych do opisu zjawisk, które nie są deterministyczne – zmiennych losowych w przypadku pojedynczych zdarzeń oraz procesów stochastycznych w przypadku zdarzeń powtarzających się (w czasie). Jako matematyczny fundament statystyki, teoria prawdopodobieństwa pełni istotną rolę w sytuacjach, w których konieczna jest analiza dużych zbiorów danych. Podstawowymi pojęciami rachunku prawdopodobieństwa są – przestrzeń zdarzeń elementarnych, z jej elementami, doświadczenie oraz zdarzenie losowe, a także prawdopodobieństwo zajścia określonego zdarzenia.

Materiał podzielony został na dwie części, stanowiące dwa odrębne tomy.

Pierwsza część składa się z trzech rozdziałów. W pierwszym i drugim przedstawiono zarys zagadnień potrzebnych do rozumienia podstaw teorii rachunku prawdopodobieństwa. Można tu znaleźć zagadnienia związane z teorią oraz wiele rozwiązanych przykładów pokazujących mechanizmy kierujące podstawowymi zasadami rachunku prawdopodobieństwa. Trzeci rozdział zawiera opis zmiennych losowych typu skokowego i ciągłego wraz z parametrami, którymi można je opisać. Wszystkie omawiane zagadnienia teoretyczne oraz ich zależności pokazane są w różnorodnych przykładach.

Druga część poświęcona została ważnym, ze względu na zastosowania w technice, rozkładom zmiennych losowych (takich jak rozkład Poissona, wykładniczy, Cauchy'ego, Laplace'a, Gaussa, Gamma i Beta oraz trapezoidalny, Rayleigha, Weibulla, Pareto czy Maxwella). Tu również oprócz opisu zamieszczono wiele rozwiązanych zadań, które ilustrują właściwości poszczególnych rozkładów zmiennych losowych. W tej części znajdują się również zestawienia najważniejszych nierówności i zależności pomiędzy omawianymi rozkładami oraz materiały uzupełniające.