

Antykruchość

O rzeczach, którym służą wstrząsy

Nassim Nicholas Taleb

Przełożyła Olga Siara



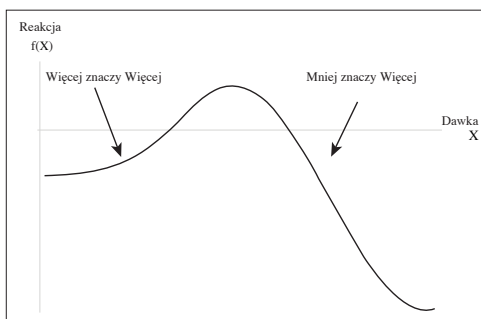
Kurhaus, Warszawa 2013

Dodatek I

Graficzny przewodnik po książce

Dla tych entuzjastów nieliniowości, którzy przedkładają rysunki nad słowa, i jedynie dla nich.

NIELINIOWOŚĆ A MNIEJ ZNACZY WIĘCEJ (& ŁOŻE PROKRUSTOWE)



Rys. 19. Rysunek ten wyjaśnia zarówno nieliniową reakcję, jak i ideę „mniej znaczy więcej”. Gdy dawka przekracza pewien poziom, korzystne efekty znikają. Jak mogliśmy zobaczyć, wszystkie zależności nieliniowe są albo wypukłe, albo wklęsłe, lub też jak na tym rysunku, mają charakter mieszany. Widać też, jak w sytuacji nieliniowości, redukcje zawodzą: łoża prokrustowe słów „to jest dla ciebie dobre” lub „to jest złe” silnie zniekształcają sens. Ponadto, rozumiemy dlaczego tak ważne są wynikające z praktycznego doświadczenia metody heurystyczne, bowiem nie doprowadzą nas one do strefy niebezpieczeństwa – słowa i opowieści mogą to sprawić. Zauważmy, że strefa „więcej znaczy wię-

we są albo wypukłe, albo wklęsłe, lub też jak na tym rysunku, mają charakter mieszany. Widać też, jak w sytuacji nieliniowości, redukcje zawodzą: łoża prokrustowe słów „to jest dla ciebie dobre” lub „to jest złe” silnie zniekształcają sens. Ponadto, rozumiemy dlaczego tak ważne są wynikające z praktycznego doświadczenia metody heurystyczne, bowiem nie doprowadzą nas one do strefy niebezpieczeństwa – słowa i opowieści mogą to sprawić. Zauważmy, że strefa „więcej znaczy wię-

cej” jest wypukła, co oznacza coraz szybciej rosnące początkowe korzyści. (W Lewantyńskim arabskim, strefa poza punktem nasycenia nazywana jest: **كترتا مثل قلتا** „więcej czegoś to jak mniej czegoś”).

Wreszcie, pokazuje to, dlaczego motywowana konkurencją skłonność do „wyrafinowania” (raczej komplikowanie rzeczy udające wyrafinowanie) jest szkodliwa, w porównaniu z wynikającym z praktycznych działań dążeniem do optymalnej prostoty.

Twierdzenie transferowe kruchości:

Zauważmy, że z twierdzenia transferowego kruchości wynika:

EKSPozycJA O KszTAŁCIE WYPUKŁYM ↔ LUBI
ZMIENNOŚĆ (DO PEWNEGO STOPNIA)

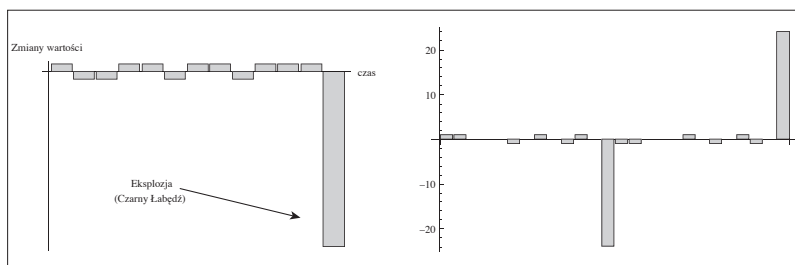
(zmienność i pozostali członkowie grona nieporządku) oraz

EKSPozycJA O KszTAŁCIE WKŁĘŚŁYM ↔ NIE LUBI
ZMIENNOŚCI

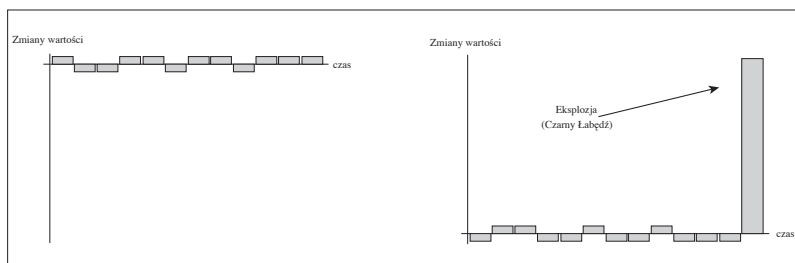
MAPOWANIE KRUCHOŚCI

W przestrzeni szeregów czasowych

Rys. 20. *Zmienność kruchości w czasie, dwa typy kruchości. Przykładowe szeregi czasowe. Oś pozioma reprezentuje czas, oś pionowa pokazuje zmienność. Te przykłady mogą dotyczyć wielu zastosowań, jak indykatory stanu zdrowia, zmiany bogactwa, twoje szczęście itd. Możemy zaobserwować małe (lub nie) pozytywne efekty i zmienność przez większość obserwowanego okresu oraz incydentalne niekorzystne efekty. Skutki niepewności mogą być bardzo dotkliwe. Zauważmy, że strata może nastąpić w dowolnym momencie i przewyższyć skumulowane do tej pory korzyści. Różnica między typem 2 (lewy wykres) a typem 1*



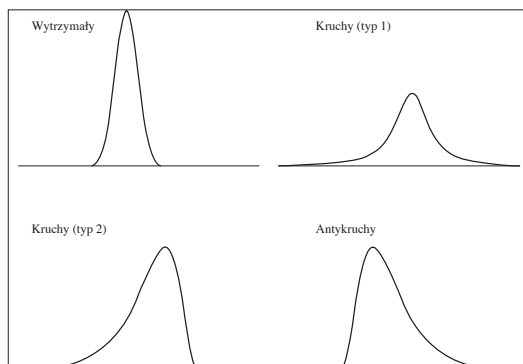
(prawy wykres) polega na tym, że w przypadku typu 2 nie występują duże efekty pozytywne, z którymi mamy do czynienia w przypadku typu 1.



Rys. 21. System po prostu wytrzymały (ale nie antykruchy) (lewy wykres): charakteryzuje go niewielka zmienność w czasie lub jej brak. Zmiany nigdy nie są znaczne. System Antykruchy (prawy wykres): niepewność przynosi dużo więcej korzyści niż strat – dokładnie przeciwnie niż w przypadku lewego wykresu na rysunku 20.

WIDZIANE W KATEGORIACH PRAWDOPODOBIEŃSTWA

Rys. 22. Oś pozioma reprezentuje wyniki, oś pionowa ich prawdopodobieństwa (tzn. rozkład prawdopodobieństwa). System wytrzymały: niewielkie korzystne i niekorzystne wyniki. System kruchy (typ 1, bardzo



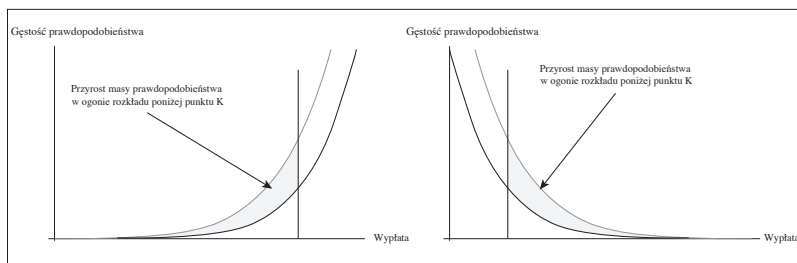
rzadki): może generować duże efekty, zarówno korzystnie, jak i niekorzystne. Dlaczego rzadko występuje? Symetrię rzadko, wręcz bardzo rzadko obserwuje się empirycznie, ale wszystkie rozkłady statystyczne

ją zakładają, dla uproszczenia. System kruchy (typ 2): obserwujemy duże, mało prawdopodobne niekorzystne efekty (często ukryte i ignorowane) i niewielkie efekty pozytywne. Istnieje możliwość zaistnienia dotkliwych efektów negatywnych (lewy wykres), w znacznie większym stopniu niż wystąpienia znacznych efektów pozytywnych, ponieważ lewa strona rozkładu jest grubsza niż strona prawa. System antykruchy: znaczne efekty pozytywne, niewielkie negatywne. Znaczne efekty pozytywne są całkiem prawdopodobne, a znaczne efekty niekorzystne dużo mniej (jeśli w ogóle możliwe). Prawy ogon rozkładu, dla efektów korzystnych, jest większy niż ogon lewy.

Tabela 9. Cztery różne efekty

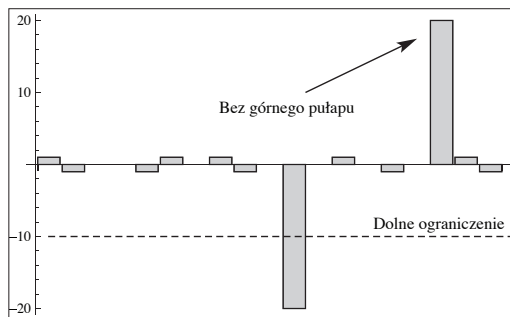
<i>Lewy ogon rozkładu</i>	<i>Prawy ogon rozkładu</i>	<i>Status systemu</i>
Cienki	Gruby	Antykruchy
Gruby	Gruby	Krchy (typ 1) (Bardzo rzadki)
Gruby	Cienki	Krchy (typ 2)
Cienki	Cienki	Wytrzymały

Kruchość charakteryzuje się grubym lewym ogonem, a więc, co istotne, jest przez to wrażliwa na zaburzenia po lewej stronie rozkładu prawdopodobieństwa.



Rys. 23. Definicja kruchości (lewy wykres): kruchość to zacieniony obszar, będący przyrostem masy prawdopodobieństwa w lewym ogonie rozkładu poniżej pewnego poziomu K zmiennej celu w reakcji na dowolną zmianę parametru dla zmiennej źródłowej – głównie „niestabilność” lub jakiś bardziej stabilny proces. Wszystkie te źródła zmienności zawrzemy w ramach parametru s_- , o którym będzie dalej mowa (gdzie udało mi się ukryć równania). Co do definicji antykruchości (prawy wykres), która nie jest do końca symetryczna, mamy lustrzane odbicie tego samego, co obserwowaliśmy poprzednio w prawym ogonie rozkładu i dodatkowo wytrzymałość w lewym ogonie. Parametrem podlegającym perturbacjom jest s_+ . Kluczowe jest to, że chociaż możemy nie być w stanie wyspecyfikować z wystarczającą dokładnością rozkładu prawdopodobieństwa, to jednak możemy zbadać reakcje systemu metodami heurystycznymi, posiłkując się „twierdzeniem o funkcji transferowej” z pracy Taleb i Douady (2012). Innymi słowy, nie ma potrzeby, byśmy znali prawdopodobieństwa przyszłych zdarzeń, ponieważ jesteśmy w stanie ocenić kruchość tych zdarzeń.

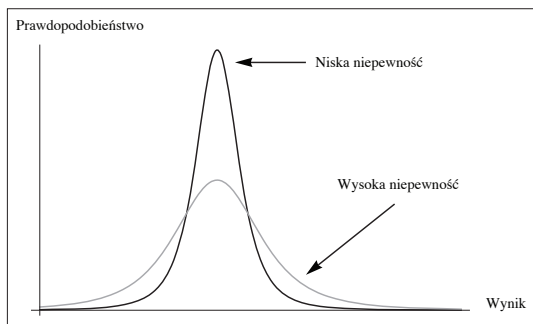
TANSFORMACJA TYPU SZTANGA W PRZESTRZENI SZEREGÓW CZASOWYCH



Rys. 24. Sztanga w przestrzeni szeregów czasowych. Ograniczenie efektów z dołu, bez górnego pułapu.

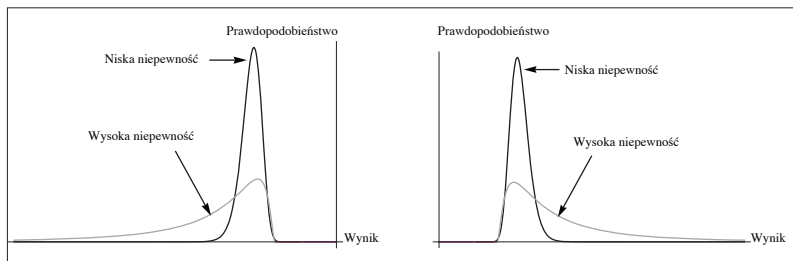
SZTANGI (TRANSFORMACJE WYPUKŁE) I ICH WŁASNOŚCI W PRZESTRZENI PROBABILISTYCZNEJ

Graficzna ilustracja koncepcji sztangi



Rys. 25. Przypadek 1 (symetria). Wprowadzenie do systemu niepewności sprawia, że przechodzimy od krzywej w kształcie dzwonu – pierwszej, z wąskim obszarem możliwych realizacji

– do drugiej krzywej, z niższym położonym wierzchołkiem, ale szerszym rozprzestawieniem. Przejście od pierwszej do drugiej krzywej powoduje przyrost pozytywnych i negatywnych niespodzianek, zarówno pozytywnych, jak i negatywnych Czarnych Łabędzi.



Rys. 26. *Przypadek 2 (lewy wykres): kruchy. Niewielkie korzyści, większe straty. Rosnąca niepewność w systemie powoduje wzrost głównie (a czasami wyłącznie) niekorzystnych wyników, po prostu negatywnych Czarnych Łabędzi. Przypadek 3 (prawy wykres): antykruchy. Rosnąca niepewność i losowość w systemie zwiększają prawdopodobieństwo bardzo korzystnych wyników i odpowiednio zwiększają oczekiwaną wypłatę. Zauważmy, że jest to dokładnie lustrzane odbicie sytuacji z lewego wykresu tego rysunku, co oznacza, że odkrycie naukowe jest, w matematycznym ujęciu, dokładnie jak przeciwieństwo opóźnienia samolotu.*

TECHNICZNE UJĘCIE PROBLEMU „TO NIE TO SAMO” GRUBEGO TONY’EGO LUB KONFLACJA ZDARZEŃ I EKSPOZYCJA NA ZDARZENIA

Uwagi te wyjaśniają również pojęcie „transformacji wypukłej”.

$f(x)$ jest ekspozycją względem zmiennej x . $f(x)$ można równoważnie nazywać „wypłatą z tytułu x ”, „ekspozycją na x ”, a nawet „użytecznością wypłaty z tytułu x ”, gdzie f definiujemy jako funkcję użyteczności, x może oznaczać cokolwiek.

Przykład: x oznacza intensywność trzęsienia ziemi, mierzoną w pewnej skali, na jakimś obszarze, $f(x)$ jest liczbą ofiar z tego powodu. Łatwo można zauważyć, że $f(x)$ może być bardziej przewidywalne niż x (jeżeli sprawimy, by ludzie nie zamieszkiwali pew-

nych obszarów lub też, by zabudowa odpowiadała pewnym standardom itd.).

Przykład: x oznacza mierzoną w metrach skalę mojego upadku w metrach z wysokości x , z której ktoś mnie zepchnął, $f(x)$ jest miarą stanu mojego zdrowia w wyniku upadku. Jasnym jest, że nie jestem w stanie przewidzieć x (tego, kto mnie zepchnie, zaś mogę $f(x)$).

Przykład: x oznacza liczbę samochodów w Nowym Jorku w jutrzejsze południe, $f(x)$ jest czasem podróży z punktu A do punktu B dla pewnej osoby. $f(x)$ może być łatwiej przewidzieć niż x (może pojechać metrem lub nawet lepiej pójść piechotą).

Niektórzy rozprawiają na temat $f(x)$, sądząc, że mówią o x . To jest problem konflacji zdarzenia i ekspozycji na nie. Ten błąd, obecny u Arystotelesa, jest w istocie powszechny w pismach filozoficznych dotyczących prawdopodobieństwa (na przykład, Hacking).

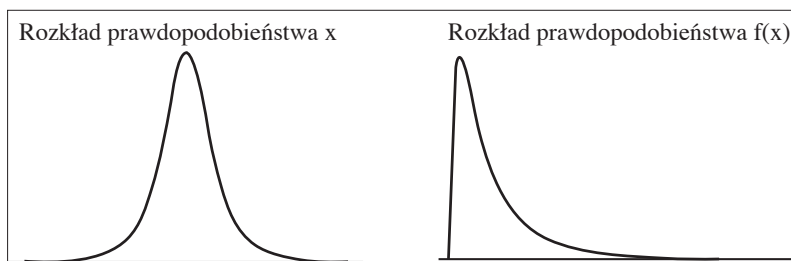
Można stać się antykruchym względem x , nie rozumiejąc x , ze względu na wypukłość $f(x)$.

Odpowiedzią na pytanie: „jak zachowujesz się w świecie, którego nie rozumiesz?” jest, po prostu: doświadczasz niepożądanych stanów $f(x)$.

Często łatwiej jest zmodyfikować $f(x)$ niż zgromadzić lepszą wiedzę na temat x . (Innymi słowy, raczej wzmacnianie wytrzymałości niż prognozowanie Czarnych Łabędzi).

Przykład: Jeśli kupię na rynku ubezpieczenie, w tym przypadku x , którego wartość spadnie o 20%, to $f(x)$ będzie niezależne od tej części rozkładu prawdopodobieństwa zmiennej x , która znajduje się poniżej 20%. (To jest przykład sztangi).

Rys. 27. Transformacja wypukła ($f(x)$ jest funkcją wypukłą względem x). Różnica między x a ekspozycją na x . Na rysunku z prawej nie



ma ryzyka spadków. Rzecz w tym, by przekształcić $f(x)$ w taki sposób by, wiedza na temat własności x po lewej stronie rozkładu stała się tak bezużyteczna, jak tylko to możliwe. Taką operację nazywamy transformacją wypukłą, której nadaliśmy tu ksywkę „sztanga”.

Problem mokrej tarcicy: Gdy ktoś bierze za $f(x)$ inną funkcję $g(x)$, która charakteryzuje się odmiennymi nieliniowościami.

Bardziej technicznie: Jeśli coś jest antykruche względem x , wtedy wariancja (lub zmienność, lub też inne miary fluktuacji) x jest korzystna dla $f(x)$, ponieważ w rozkładach, które są skośne, średnia zależy od wariancji, a także, jeśli rozkład jest prawoskośny, wartość oczekiwana rośnie wraz z wariancją (na przykład, w przypadku rozkładu lognormalnego, wartość średnia zawiera składnik $+1/2 \sigma^2$).

Co więcej, rozkład prawdopodobieństwa dla $f(x)$ jest istotnie różny, niż w przypadku x , szczególnie gdy występują nieliniowości.

Gdy $f(x)$ jest wypukła (wklęsła) monotonicznie, $f(x)$ jest prawoskośna (lewoskośna).

Gdy $f(x)$ jest rosnąca oraz wypukła z lewej i wklęsła po prawej stronie, wówczas rozkład prawdopodobieństwa $f(x)$ charakteryzuje się cieńszym ogonem niż rozkład x . Na przykład, w teorii perspektywy Kahnemana – Tversky’ego, tak zwana użyteczność zmian bogactwa jest bardziej „wytrzymała” niż użyteczność samego bogactwa.

Dlaczego wypłaty mają większe znaczenie niż prawdopodobieństwo (technicznie): Gdy $p(x)$ jest funkcją gęstości prawdopodobieństwa, wartość oczekiwana, to jest $\int f(x) p(x) dx$, zależeć będzie raczej od f , a nie od p , a ponadto im bardziej nieliniowa f , tym silniejsza będzie to zależność od f , raczej niż od p .

CZWARTA ĆWIARTKA (TALEB, 2009)

Idea polega na tym, że zdarzenia rzadkie są nie do policzenia (w domenie rozkładów gruboogonowych), ale możemy ocenić naszą ekspozycję na nie.

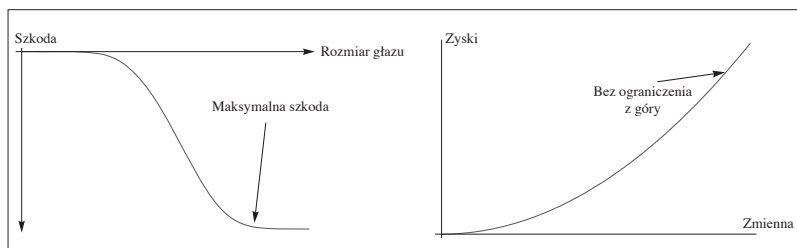
Tabela 10.

	<i>LEKKOOGONOWY ROZKŁAD WZGLĘDEM X</i>	<i>GRUBOOGONOWY ROZKŁAD WZGLĘDEM X</i>
$f(x)$ „łagodzi” poprzez odcięcie skrajnych efektów, tj. przypadek wypukło - wklęsły	Wynik bardzo wytrzymały	Wynik dość wytrzymały
$f(x)$ wklęsło-wypukła, wzmacnia najbardziej odległe efekty	Wynik wytrzymały (tak trochę)	CZWARTA ĆWIARTKA Kruchy (jeżeli $f(x)$ jest wklęsła) lub antykruchy

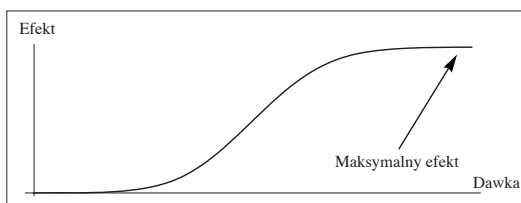
WYPUKŁOŚCI LOKALNE I GLOBALNE (TECHNICZNE)

Nic nie jest nieograniczone w naturze – śmierć jest maksymalnym ostatecznym rezultatem dla jednostki. Więc wszystko kończy się w strefie wypukłości na jednym końcu, w strefie wklęsłości na drugim.

W istocie, dla istot biologicznych istnieje punkt maksymalnej straty. Powróćmy do wklęsłego obrazu głazu i kamyków z Rozdziału 18: rozszerzając analizowany przedział dostrzegamy, że ograniczoność straty prowadzi do wypukłości w jakimś obszarze.



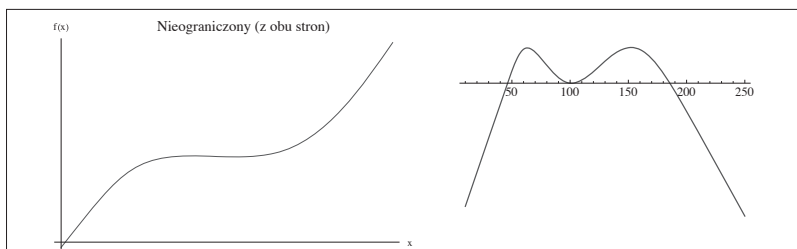
Rys. 28. Lewy wykres ilustruje opowieść o głazie i kamykach z Rozdziału 18, w przypadku rozszerzonego przedziału. W pewnym momencie, wklęsłość przechodzi w wypukłość, gdy straty. Prawy wykres ilustruje silną antykruchość, bez ograniczenia z góry (co prowadzi do Ekstremist stanu). Takie efekty są możliwe jedynie w przypadku zmiennych ekonomicznych, jak na przykład sprzedaż książek, lub też w przypadku zjawisk nieograniczonych lub niemal nieograniczonych. Nie potrafię wskazać tego rodzaju efektu w naturze.



Rys. 29. Słaba antykruchość (Przeciętno stan), ze skończonym maksimum. Zjawisko typowe w naturze.

NIETYPOWE NIELINIOWOŚCI (BARDZO TECHNICZNE)

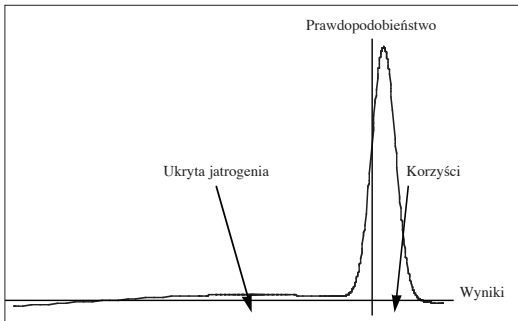
Następne dwa rodzaje nieliniowości są niemal niespotykane poza obszarem zmiennych ekonomicznych; są one w szczególności ograniczone do zjawisk powodowanych przez instrumenty pochodne.



Rys. 30. Wykres lewy obrazuje wypukło-wklęsłą funkcję rosnącą, przeciwnieństwo ograniczonych zależności typu dawka – efekt, które obserwujemy w naturze. To prowadzi nas do typu 2, kruchy (bardzo, ale to bardzo grube ogony rozkładu). Wykres prawy ilustruje przypadek najbardziej niebezpieczny ze wszystkich: pseudowypukłość. Lokalna antykruchość, globalna kruchość.

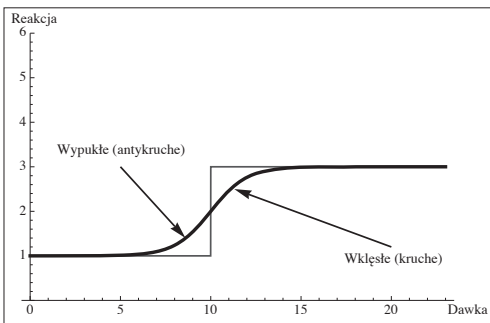
NIELINIOWOŚCI MEDYCZNE I ICH ODPOWIEDNIKI PROBABILISTYCZNE (ROZDZIAŁY 21 & 22)

Rys. 31. Jatrogenia w medycynie: przypadek niewielkich korzyści i wielkich strat typu Czarny Łabędź, przedstawiony w przestrzeni probabilistycznej. Jatrogenia występuje wtedy, gdy mamy do czynienia z niewielkimi, dającymi się zidentyfikować korzyściami (na przykład, uniknięcie nieznacznego dyskomfortu lub drobnej infekcji), ale jesteśmy narażeni na Czarne Łabędzie z odłożonymi i ukrytymi olbrzymimi skutkami ubocznymi (jak śmierć). Te rozłożone według wklęsłej krzywej korzyści ze stosowania medycyny są dokładnie takie, jak ze sprzedaży opcji fi-



nansowych (olbrzymie ryzyko) dla zapewnienia znikomych natychmiastowych zysków, z uzasadnieniem, że „nie stwierdzono negatywnych skutków” [ang. evidence of no harm – przyp. tłum.].

Mówiąc w skrócie, dla osoby zdrowej istnieje niewielkie prawdopodobieństwo fatalnych skutków (pomijanych, ponieważ nie są widoczne, więc nie są brane pod uwagę) i znaczne prawdopodobieństwo umiarkowanych korzyści.

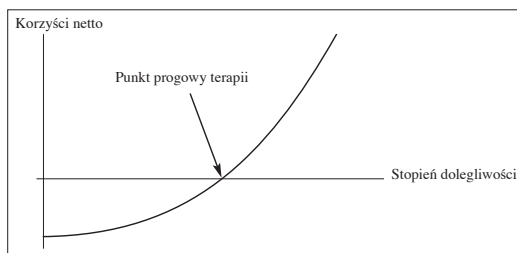


Rys. 32. Nieliniowości w biologii. Wypukłowo-wklęsły kształt niewątpliwie pochodzi od dowolnego procesu wzrostu (monotonicznego, tj. bez spadków), który jest ograniczony, z wartościami maksimum i minimum,

tj. proces nie zmierza do nieskończoności z żadnej strony. Na niskich poziomach, zależność dawka – efekt jest wypukła (stopniowy wzrost efektywności). Kolejne dawki stają się coraz mniej efektywne lub zaczynają szkodzić. Podobny schemat dotyczy konsumpcji jakiegokolwiek dobra z nadmierną częstością. Ten wykres bez wątpienia dotyczy dowolnego procesu ograniczonego po obu stronach, ze znanymi wartościami minimum i maksimum (stan nasycenia), co dotyczy też szczęścia.

Na przykład, jeśli ktoś uważa, że istnieje maksymalny poziom szczęścia i nieszczęścia, to ogólny kształt krzywej, wypukłej z lewej strony i wklęsłej po prawej, będzie z konieczności odnosił się do szczęścia

(wystarczy zamienić „dawkę” na „bogactwo” oraz „reakcję” na „szczęście”). Teoria perspektywy Kahnemana-Tversky’ego modeluje podobny kształt dla krzywej „użyteczności” zmian bogactwa, co autorzy stwierdzili empirycznie.

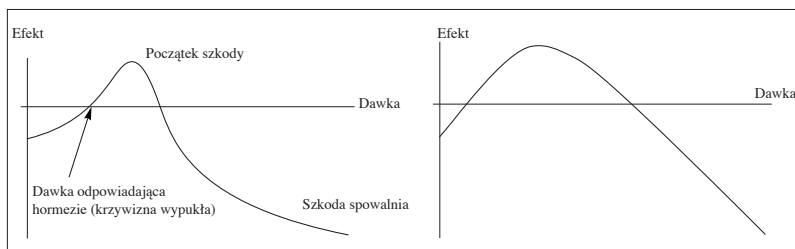


Rys. 33. Przypomnijmy sobie przykład z nadciśnieniem tętniczym. Na osi pionowej odkładamy korzyści z leczenia, a na osi poziomej – stopień dolegliwości.

Strzałka wskazuje na poziom, przy którym probabilistycznie mierzone korzyści zrównują się z probabilistycznie mierzonymi stratami. Jatrogenia znika nieliniowo jako funkcja stopnia dolegliwości. Wynika stąd, że kiedy pacjent jest bardzo chory, rozkład przesuwają się w stronę antykruchości (grubszy prawy ogon), z czym wiążą się znaczne korzyści z leczenia względem możliwej jatrogenii, a więc niewiele do stracenia.

Zauważmy, że jeśli zwiększymy intensywność leczenia, znajdziemy się w obszarze wklęsłości krzywej wraz z osiągnięciem maksymalnych korzyści, czego nie obejmuje rysunek – gdyby więc spojrzeć szerzej na ten proces, otrzymalibyśmy rysunek zbliżony do poprzedniego.

Rys. 34. Lewy wykres ilustruje zjawisko hormezy organizmu żywego (podobnie jak na rysunku TK): obserwujemy fazę korzyści, gdy dawka rośnie (początkowo według krzywizny wypukłej), po czym korzystny efekt słabnie przechodząc w fazę straty, kiedy zwiększamy dawkę trochę bardziej (początkowo według krzywizny wklęsłej); następnie obserwujemy, jak krzywa wypłaszcza się na poziomie maksymalnej straty (powyżej pewnego punktu, organizm staje się martwy, co pokazuje, że dla biologii właściwy jest przypadek ograniczoności procesu, który jest najgorszym możliwym scenariuszem. Na prawym wykresie, widzimy nie-



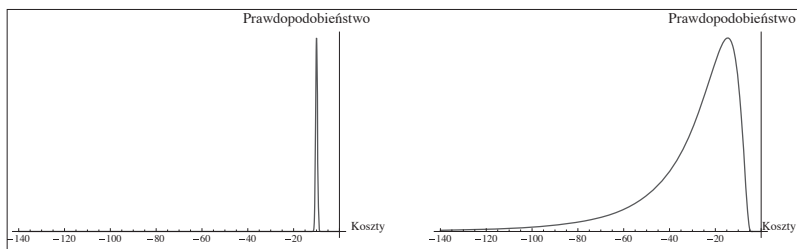
poprawny przebieg hormezy z podręczników medycyny, z początkową wklęsłością, zaczynającą się od fazy liniowej lub lekko wklęsłej.

ODWROTNOŚĆ PROBLEMU INDYKA

Rys. 35. Antykruchy system, odwrotność problemu indyka: niewidoczne rzadkie zdarzenie ma charakter pozytywny. Gdy patrzycie na dodatnio skośny (antykruchy) szereg czasowy i wnioskuje o tym, czego nie widać, pomijacie to, co dobre i nie doszacowujecie korzyści (Pisano, 2006a, 2006b, błędne ujęcie). Na prawym rysunku, inny problem harwardzki, podniesiony przez Froota (2001). Wypełniony obszar odpowiada temu, czego skłonni jesteśmy nie dostrzegać w małych próbach, z racji niewystarczającej liczby obserwacji. Ciekawe, że zaszarzony obszar rośnie wraz z błędem modelu. W bardziej technicznych partiach Dodatku obszar ten oznaczony będzie ωB (indyk) i ωC (odwrotność problemu indyka).

RÓŻNICA MIĘDZY ESTYMATORAMI PUNKTOWYMI A ROZKŁADAMI

Zastosujmy tę metodę analizy do prześledzenia, jak planiści popełniają błędy, które popełniają i dlaczego deficyty zwykle są większe od planowanych.



Rys. 36. Rozbieżność między prognozami a rzeczywistością: rozkład prawdopodobieństwa kosztów projektów w głowach planistów (lewy wykres) i w rzeczywistości (prawy wykres). Zgodnie z pierwszym wykresem, zakładają oni, że koszty będą zarówno niskie, jak i całkiem pewne. Prawy wykres pokazuje wyniki, które są zarówno gorsze, jak i bardziej rozproszone, a w szczególności te gorsze wyniki są bardziej prawdopodobne. Zauważmy, że kruchość wzrasta z tytułu pęcznienia lewego ogona rozkładu.

Niezrozumienie skutków niepewności dotyczy deficytów rządowych, planów zawierających elementy IT, czasu podróży (w mniejszym stopniu) i jeszcze wielu innych. Wykorzystamy jeszcze dalej ten rysunek by zilustrować efekt błędu modelu, wynikający z niedoszacowania kruchości, jeśli parametr założony jako stały jest losowy. A to właśnie jest plagą ekonomii w wydaniu biurokratów (patrz, dalsze rozważania).

Tam gdzie większość modeli ekonomicznych kruszeje i wysadza nas w powietrze

Gdy używałem terminu „techniczny” w głównym tekście, trochę oszukiwałem. Ale już nie tutaj.

Niespójność Markowitza: załóżmy, że ktoś powiedział wam, że prawdopodobieństwo pewnego zdarzenia wynosi dokładnie zero. Pytacie go, skąd to wie. „Wiem to od Baala”, brzmi odpowiedź. W tym przypadku osoba ta jest spójna, ale będzie uważana za niewiarygodną przez tych, którzy nie są wyznawcami Baala. Z drugiej strony, gdyby osoba ta powiedziała wam: „Oszacowałem to prawdopodobieństwo na zero”, to mamy problem. Osoba ta jest bowiem zarówno niewiarygodna, jak i niespójna. Coś, co zostało oszacowane, musi mieć błąd szacunku. Prawdopodobieństwo nie może więc równać się zeru, jeśli wynika z estymacji, a dolne ograniczenie szacunku zależy od błędu estymacji: im większy błąd estymacji, tym większe prawdopodobieństwo, ale tylko do pewnego poziomu. Jak wynika z argumentacji Laplace’a w kwestii całkowitej niewiedzy, wraz ze wzrostem błędu estymacji do nieskończoności, prawdopodobieństwo zmierza do $\frac{1}{2}$.

Powróćmy do konsekwencji błędnego myślenia; przyjmijmy teraz, że jakkolwiek estymacja parametru, a następnie użycie tej wartości w jakimś równaniu, jest różne od jednoczesnej estymacji wszystkich parametrów równania (tak samo, jak to było w przypad-

ku zdrowia babci, przeciętnej temperatury, więc termin »estymacja« nie jest istotny, a to, na czym nam zależy, to wiedza na temat przeciętnego stanu zdrowia przy różnych wartościach temperatury). No i Markowicz wykazał się niespójnością w swoim rozumowaniu, rozpoczynając swój „wiekopomny” artykuł: „Załóżmy, że znamy E i V ” (czyli wartość oczekiwaną i wariancję). Na końcu artykułu przyznaje on, że muszą być one oszacowane oraz, co gorsza, z zastosowaniem zestawu technik statystycznych i „zdrowego rozsądku praktyków”. Otóż, jeśli parametry te trzeba oszacować, co pociąga za sobą błąd estymacji, to całe rozumowanie powinno być zapisane inaczej i rzecz jasna nie byłoby artykułu – i nie tylko artykułu Markowicza, ale nie byłoby kryzysów, nowoczesnych finansów, wyznawców kruchości uczących studentów bezwartościowej wiedzy...

Prosty przykład: deficyty budżetowe

Posługujemy się przykładem deficytów budżetowych ze względu na to, że obecnie stosowane przez rządy i agencje rządowe sposoby ich obliczania pomijają czynnik wypukłości (i mają one problem, by to zaakceptować). W rzeczywistości, instytucje te nie biorą tego czynnika pod uwagę. Przykład, który to ilustruje:

(a) Zaniedbanie stochastycznej natury zmiennej, co do której wiemy, że wpływa na zachowanie modelu, ale która wydaje się deterministyczna (i ma wartość ustaloną), oraz

(b) F , funkcja tej zmiennej, jest wypukła lub wklęsła względem tej zmiennej.

Powiedzmy, że rząd szacuje stopę bezrobocia w ciągu trzech najbliższych lat na średnio 9%; używa swoich modeli ekonometrycznych, by wyprognozować deficyt budżetowy B na poziomie 200 miliardów w lokalnej walucie. Ale nie bierze pod uwagę (co tyczy się niemal wszystkiego w ekonomii), że bezrobocie jest zmienną sto-

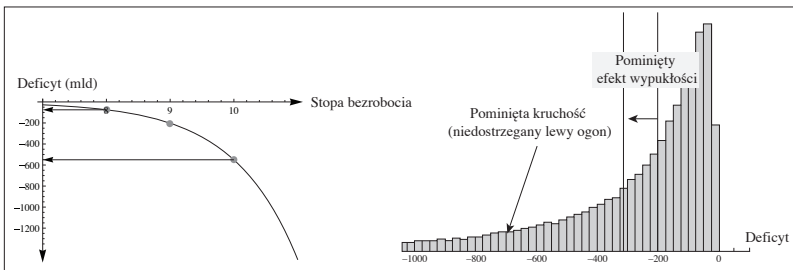
chastyczną. W rzeczywistości, stopa bezrobocia wahała się w ciągu trzech lat średnio o 1%. Możemy następująco policzyć efekt błędu:

Stopa bezrobocia równa 8%, deficyt $B(8\%) = -75$ mld (poprawa o 125 mld)

Stopa bezrobocia równa 9%, deficyt $B(9\%) = -200$ mld

Stopa bezrobocia równa 10%, deficyt $B(10\%) = -550$ mld (pogorszenie o 350 mld)

Obciążenie z tytułu wklęsłości, lub inaczej ujemne obciążenie z tytułu wypukłości, wynikające z niedoszacowania deficytu wynosi -112,5 mld, ponieważ $\frac{1}{2} \{B(8\%) + B(10\%)\} = -312$ mld, nie zaś -200 mld. To jest właśnie przykład odwrotności kamienia filozoficznego [ang. *the inverse philosopher's Stone* – przyp. tłum.].



Rys. 37. Transformacje nieliniowe pozwalają na wykrycie zarówno obciążenia z tytułu wypukłości modelu, jak i kruchości. Graficzna ilustracja przykładu: histogram wyników symulacji Monte Carlo deficytu budżetowego jako lewoskośnej zmiennej losowej, naturalnie w wyniku randomizacji stopy bezrobocia, względem której deficyt jest funkcją wklęsłą. Metoda punktowej estymacji zakładałaby zachowanie zgodne z rozkładem typu delta Diraca w punkcie -200, co pociąga za sobą niedoszacowanie zarówno oczekiwanego deficytu (-312), jak i jego ogonowej kruchości.

**ZASTOSOWANIE: MODEL RICARDO I LEWY OGON
– CENA WINA MOŻE SIĘ ZMIENIAĆ**

Od niemal 200 lat dyskutujemy nad pojęciem zwanym „przewaga komparatywna”, wprowadzonym przez ekonomistę Davida Ricardo. W skrócie, oznacza ono, że kraj powinien prowadzić pewną politykę, na podstawie przewag komparatywnych w odniesieniu do wina lub sukna. Załóżmy, że dany kraj jest dobry w produkcji zarówno wina, jak i sukna, i jest w tym lepszy od swoich sąsiadów, z którymi może prowadzić wolny handel. W takiej sytuacji rzecz jasna optymalną strategią byłoby specjalizowanie się w tym, spośród wina i sukna, co odpowiada danemu krajowi najbardziej i minimalizuje koszty alternatywne. Wszyscy byłiby wtedy zadowoleni. Analogicznie, według ekonomisty Paula Samuelsona, jeśli ktoś jest najlepszym lekarzem w mieście i równocześnie najlepszą sekretarką, to najlepiej byłoby, gdyby owa osoba została dobrze zarabiającym lekarzem – ponieważ minimalizowałoby to koszty alternatywne – pozwalając komuś innemu być sekretarką i kupować od tej osoby usługi sekretarki.

Zgadzam się z tym, że są pewne korzyści z jakiejś formy specjalizacji, ale nie wynika z modeli, które te korzyści uzasadniają. W rozumowaniu tym występuje następujący mankament. To prawda, byłoby niewyobrażalne, by doktor został sekretarką na część etatu tylko dlatego, że jest w tym dobry. Ale jednocześnie możemy niezawodnie założyć, że bycie lekarzem gwarantuje pewną stabilizację zawodową: ludzie zawsze będą chorować, a także z profesją lekarza wiąże się wyższy status społeczny niż z bycia sekretarką, co czyni profesję lekarza bardziej atrakcyjną. Ale teraz załóżmy, że w świecie składającym się z dwóch krajów, jeden z krajów specjalizuje się w produkcji wina, z zamiarem sprzedaży swojego specjału do drugiego kraju, a *tymczasem nagle cena wina spada gwałtownie*. Jakaś zmian w gustach konsumentów spowodowała zmianę ceny. Analiza Ricardo zakłada, że zarówno cena rynkowa wina, jak i koszty jego produkcji pozostają stałe i że historia ta nie ma „drugiego dna”.

Tabela 11. *ORYGINALNY PRZYKŁAD RICARDO*
(JEDNOSTKOWE KOSZTY PRODUKCJI)

	SUKNO	WINO
Brytania	100	110
Portugalia	90	80

Rozumowanie: tabela powyżej pokazuje koszty produkcji, znormalizowane względem ceny sprzedaży jednostki każdego dobra, to znaczy zakładając, że dobra te są wymieniane po równych cenach (jednostka sukna za jednostkę wina). Jako paradoks wydaje się sytuacja, w której Portugalia produkuje sukno taniej niż Brytania, ale powinna je stamtąd kupować, wykorzystując do tego przychody ze sprzedaży wina. Przy braku kosztów transakcyjnych i transportu, efektywnym jest, by Brytania produkowała jedynie sukno, a Portugalia tylko wino.

Koncepcja ta zawsze przyciągała ekonomistów ze względu na swój paradoksalny i kontrintuicyjny aspekt. Na przykład, w artykule „Why Intellectuals Don’t Understand Comparative Advantage” (Krugman, 1998), Paul Krugman, który ma sam problem ze zrozumieniem tego pojęcia, ponieważ tekst ten i jego techniczna część pokazują, że jest on jest całkowicie pozbawiony wiedzy na temat rzadkich zdarzeń i zarządzania ryzykiem, a jednocześnie kpi z innych intelektualistów, takich jak S. J. Gould, którzy rozumieją ideę rzadkich zdarzeń, chociaż raczej intuicyjnie niż analitycznie. (Jasnym jest, że nie można mówić o zwrotach i zyskach bez dyskontowania tych korzyści równoważącymi je ryzykami). Artykuł ten pokazuje, że Krugman popełnia istotny i niebezpieczny błąd mylenia funkcji średniej ze średnią funkcji.

Założmy teraz, że ceny wina i sukna są zmienne – czego Ricardo nie zakładał – i że liczby powyższe liczby odzwierciedlają ich nieobciążone długookresowe wartości. Dalej założmy, że ceny zachowują się zgodnie z rozkładem gruboogonowym. Lub też za-

łożmy, że ich koszty produkcji zmieniają się zgodnie z rozkładem gruboogonowym.

Jeśli cena wina na rynkach międzynarodowych wzrośnie o 40%, wtedy korzyści będą oczywiste. Gdyby jednak cena spadła w jednakowym stopniu, o 40%, przyniosłoby to ogromne straty, znacznie większe niż korzyści przy tym samym, co do skali, wzroście cen. Ekspozycję charakteryzującą w tym przypadku wklęsłości – silne wklęsłości.

Jasnym więc jest, że gdyby cena spadła o 90%, to efekt byłby katastrofalny. Po prostu wyobraźcie sobie, co stałoby się z waszym gospodarstwem domowym, gdyby dotknęła je nagle i nieprzewidziana obniżka wynagrodzenia o 40%. W rzeczy samej, obserwowaliśmy na przestrzeni historii problemy krajów, które specjalizowały się produkcji pewnych towarów, surowców i płodów rolnych, które okazywały się być nie dość, że niestabilne, to skrajnie niestabilne. A katastrofa niekoniecznie musi być wynikiem zmienności cen, ale może być rezultatem problemów w produkcji: produkcja rolna może ustać z powodu zarazków, złej pogody lub jakiejś innej przeszkody.

Zły urodzaj, jak ten, który był przyczyną wywołanego brakiem ziemniaków głodu w Irlandii w ciągu dekadę około 1850 roku, spowodował śmierć miliona ludzi i emigrację kolejnego miliona (cała ludność Irlandii, w momencie pisania tej książki, wynosi około 6 milionów, jeśli uwzględnimy część północną). Bardzo trudno jest gruntownie odmienić posiadane zasoby – nie tak, jak w przykładzie lekarza i sekretarki, kraje nie mają takich możliwości zmiany. Rzeczywiście, monokultura (skoncentrowanie się na jednej uprawie) okazała się zabójcza w historii – jeden zły urodzaj powoduje wyniszczający głód.

Jeszcze jedna różnica w stosunku do przykładu lekarza i sekretarki polega na tym, że kraje nie mają rodziny ani przyjaciół. Lekarz może liczyć na wsparcie ze strony społeczności, kręgu przyjaciół, kolektywu, który dba o niego, teścia, od którego może wziąć pożyczkę w przypadku, gdyby chciał zmienić profesję, państwa nad sobą,

które w razie czego udzieli pomocy. Kraje nie mają tego wszystkiego. Co więcej, lekarz ma oszczędności; kraje są raczej dłużnikami.

Znowu więc napotykamy na kruchość ze względu na efekty drugiego rzędu.

Jak działa specjalizacja: Czytelnik nie powinien interpretować tego co mówię w ten sposób, że specjalizacja nie jest dobra – mówię jedynie tyle, że powinno się określać specjalizację, biorąc pod uwagę kwestie kruchości i efektów drugiego rzędu. Obecnie sędzę, że Ricardo ma absolutnie rację, ale nie wynika to z zaprezentowanych przez niego modeli. W sposób organiczny, systemy pozbawione odgórnego sterowania podlegałyby specjalizacji stopniowo, powoli i w długim okresie, metodą prób i błędów, aż do osiągnięcia właściwego poziomu specjalizacji – a nie przez to, że jakiś biurokrata stosowałby model. Powtórzmy, systemy robią małe błędy, projektowanie popełnia duże.

Stąd, decyzje społecznego planisty ściśle oparte na wyrastającym z intuicji modelu Ricardo doprowadziłyby do katastrofy; pogodzenie się z powolnym procesem eksperymentowania efektywność przyniosłoby efektywność – prawdziwą efektywność.

BARDZIEJ OGÓLNA METODOLOGIA WYKRYWANIA BŁĘDÓW MODELU

Efekty drugiego rzędu w modelu i kruchość: Załóżmy, że mamy właściwy model (co jest bardzo daleko idącym założeniem), ale nie mamy pewności, co do wartości parametrów. Uogólniając przykład z deficytem/bezrobociem, użyty w poprzednim punkcie, przyjmijmy, że posługujemy się funkcją f , zdefiniowaną w prosty sposób: $f(x \mid \bar{\alpha})$, gdzie zakłada się, że $\bar{\alpha}$ jest wartością oczekiwaną zmiennej wejściowej i gdzie przyjmujemy, że ϕ jest rozkładem zmiennej α w dziedzinie $\mathcal{P}_\alpha$, $\alpha = \int_{\mathcal{P}_\alpha} \alpha \varphi(\alpha) d\alpha$.

Kamień filozoficzny: Już samo to, że α ma charakter losowy (podlega estymacji) mogłoby doprowadzić do obciążenia wyniku, gdybyśmy zadali zakłócenie *od wewnątrz* (całki), tzn. uczynili stochastycznym parametr uważany do tej pory za stały. Odpowiednio, obciążenie z tytułu wypukłości może być łatwo zmierzone jako różnica między (a) wartością funkcji f scałkowanej po wszystkich wartościach α , (b) wartością f dla pojedynczego α przyjętego jako średnia jego wartość. Obciążenie z tytułu wypukłości (kamień filozoficzny) ω_A wynosi więc:

$$\omega_A \equiv \int_{\varphi_x} \int_{\varphi_\alpha} f(x|\alpha) \varphi(\alpha) d\alpha dx - \int_{\varphi_x} f(x) \left(\int_{\varphi_\alpha} \alpha \varphi(\alpha) d\alpha \right) dx$$

Główne równanie: Kruchosc jest jedynie częściowym kamieniem filozoficznym poniżej punktu K , dlatego brakująca kruchosc ω_B jest wyznaczona przez porównanie wartości dwóch całek poniżej punktu K , ażeby uchwycić oddziaływanie na lewy ogon rozkładu:

$$\omega_B(K) \equiv \int_{-\infty}^K \int_{\varphi_\alpha} f(x|\alpha) \varphi(\alpha) d\alpha dx - \int_{-\infty}^K f(x) \left(\int_{\varphi_\alpha} \alpha \varphi(\alpha) d\alpha \right) dx$$

co może być aproksymowane w drodze interpolacji, z wykorzystaniem dwóch wartości α , oddzielonych od punktu środkowego o $\Delta\alpha$, średnie odchylenie α , a następnie wyznaczając:

$$\omega_B(K) \equiv \int_{-\infty}^K \frac{1}{2} (f(x|\bar{\alpha} + \Delta\alpha) + f(x|\bar{\alpha} - \alpha)) dx - \int_{-\infty}^K f(x|\bar{\alpha}) dx$$

Zauważmy, że antykruchosc ω_c jest wyznaczana poprzez całkowanie od K do nieskończoności. Możemy badać ω_B , przeprowadzając estymację punktową wartości funkcji f dla poziomów $X \leq K$

$$\omega'_B(X) = \frac{1}{2} (f(X|\bar{\alpha} + \Delta\alpha) + f(X|\bar{\alpha} - \alpha)) - f(X|\bar{\alpha})$$

a stąd

$$\omega_B(K) = \int_{-\infty}^K \omega'_B(x) dx$$

co prowadzi nas do heurystyki wykrywania kruchości (Taleb, Canetti, et al., 2012). W szczególności, jeśli założymy, że $\omega'_B(X)$ ma stały znak dla $X \leq K$, wtedy $\omega_B(K)$ jest tego samego znaku. Heurystyka wykrywania kruchości polega na stosowaniu perturbacji w ogonach rozkładu do testowania kruchości, poprzez sprawdzanie wartości funkcji $\omega'_B(X)$ dla dowolnego poziomu X .

Tabela 12.

MODEL	ŹRÓDŁO KRUCHOŚCI	REMEDIU
Teoria portfelowa, analiza średnio-wariacyjna itp.	Zakłada znajomość wartości parametrów, nie całkuje modeli po wszystkich możliwych wartościach parametrów, opiera się na (bardzo niestabilnych) korelacjach. Zakłada, że ω_A (obciążenie) i ω_B (kruchość) = 0	1/n (rozkładaj swoje ekspozycje na tak dużą ich liczbę, jak to możliwe, sztangi, stopniowy i organiczny proces budowania itp)
Teoria przewag komparatywnych Ricardo	Pominięcie losowości w przypadku ceny wina może powodować całkowite odwrócenie alokacji. Zakłada, że ω_A (obciążenie) i ω_B (kruchość) = 0	Systemy naturalne znajdują własne wzorce alokacji metodą eksperymentowania
Optymalizacja według Samuelsona	Źródła losowości skoncentrowane we wklęsłości funkcji straty. Zakłada, że ω_A (obciążenie) i ω_B (kruchość) = 0	Szersze uwzględnienie losowości

Model Arrow – Debreu, kraty, modele przestrzenie stanów	Komiczny błąd: zakłada się doskonałą znajomość wszystkich wyników i prawdopodobieństw.	Wprowadzenie meta-prawdopodobieństw zmienia całkowicie wskazania modelu
Modele dywidendy, przepływów pieniężnych	Pomijanie stochastyczności powoduje efekty wypukłości. Zwykle zakładają ω_C (antykruchosć) = 0	Heurystyka

Błędy teorii portfelowej: Zwróćmy uwagę na błędne rozumowanie wyznawców Markowitza: teoria portfelowa kusi ludzi dywersyfikacją, bo to lepsze niż nic. Otóż nie tak, wy finansowi głupcy: skłania ich do optymalizacji, a więc do nadmiernej alokacji. Nie nakłania ona ludzi do mniej ryzykownych zachowań poprzez dywersyfikację, ale powoduje, że zajmują bardziej otwarte pozycje, sądząc, że są zabezpieczeni statystycznymi własnościami modelu – to czyni ich wrażliwymi na błąd modelu, zaś szczególnie wrażliwymi na niedoszacowanie rzadkich zdarzeń. By przekonać się, jak bardzo, rozważmy dwóch inwestorów, stojących w obliczu wyboru alokacji, spośród trzech pozycji: gotówka oraz papiery wartościowe A i B . Inwestor, który nie zna własności statystycznych papierów A i B , i wie, że ich nie zna, zainwestuje – powiedzmy – część, której nie chce stracić w gotówkę, a resztę w A i B – zgodnie z jakąkolwiek będącą tradycyjnie w użyciu wiedzą praktyczną. Inwestor, który uważa, że zna owe własności statystyczne, zawarte w parametrach σ_A , σ_B , ρ_{AB} , dokona alokacji ω_A , ω_B w taki sposób, żeby osiągnąć łączne ryzyko na pewnym docelowym poziomie (pomińmy oczekiwany zwrot z tej alokacji). Im niższe jego postrzeganie korelacji ρ_{AB} , tym gorsza jego ekspozycja na ryzyko modelu. Zakładając, że uważa on korelację ρ_{AB} za równą 0, dokona on alokacji nadmiernej o $1/3$ dla skrajnych zdarzeń. Ale jeśli ten biedny inwestor kieruje się iluzją, że ta korelacja wynosi -1 , zrealizuje nadmierną alokację w maksy-

malnym stopniu względem aktywów A i B . Jeśli ten inwestor posługuje się dźwignią finansową, dochodzimy do historii Long Term Capital Management, który dał się zmylić wartościami parametrów. (W rzeczywistości, w odróżnieniu od prac ekonomicznych, wszystko może się zmieniać; i zmienia się, na Baal!). Możemy powtórzyć to rozumowania dla każdego parametru σ i przekonać się, jak obniżona percepcja wartości tego σ prowadzi do nadmiernej alokacji.

Zauważyłem to jako *trader* – i stało się to moją obsesją – że korelacje nigdy nie były takie same dla różnych pomiarów. Powiedzenie, że były niestabilne, to za mało: 0,8 na przestrzeni długiego okresu staje się – 0,2 w kolejnym długim okresie. Frajerska gra w czystej postaci. W czasach napięć na rynkach, korelacje podlegają nawet bardziej gwałtownym zmianom – bez jakiegokolwiek godnej zaufania regularności, wbrew próbom modelowania „korelacji w okresie napięć”. Taleb (1997) zajmuje się efektem stochastycznych korelacji: bezpieczną strategią jest skracanie pozycji przy korelacji na poziomie 1 i ponowne kupowanie przy poziomie – 1 – to z grubsza odpowiada temu, jak działa heurystyczna metoda $1/n$.

Kryterium Kelly’ego a Markowitz: Ażeby zastosować w pełni optymalizację à la Markowitz, trzeba w całości znać łączny rozkład prawdopodobieństwa wszystkich aktywów w przyszłości, a dodatkowo dokładną postać funkcji użyteczności dla bogactwa dla wszystkich przyszłych okresów. I to bez żadnych błędów! (Widzieliśmy, że błędy estymacji powodują eksplozję systemu). Metoda Kelly’ego, opracowana w mniej więcej tym samym okresie, nie wymaga znajomości łącznego rozkładu ani funkcji użyteczności. W praktyce, to co niezbędne, to relacja oczekiwanych zysków do najgorszego akceptowanego zwrotu – dostosowywane dynamicznie, dla uniknięcia całkowitej ruiny. W przypadku transformacji typ sztangy, najgorszy akceptowany zwrot jest zagwarantowany. I błąd modelu jest znacznie, ale to znacznie bardziej umiarkowany w przypadku kryterium Kelly’ego. Thorp (1971, 1998), Haigh (2000).

Godny respektu Aaron Brown utrzymuje, że koncepcje Kelly'ego zostały odrzucone przez ekonomistów – mimo ich praktycznej atrakcyjności – ponieważ uwielbiają oni ogólne teorie, obejmujące ceny wszystkich aktywów.

Zauważmy, że podejście ograniczonych prób i błędów jest zgodne z kryterium Kelly'ego, kiedy ma się wyobrażenie, co do potencjalnego zwrotu – ale nawet, jeśli ktoś nie ma pojęcia co do zwrotu, jeśli tylko straty są ograniczone, efekt będzie odporny i metoda powinna dostarczyć lepszych wyników niż podejście wyznawcy kruchości, Markowitza.

Finansowanie przedsiębiorstw [ang. *corporate finance* – przyp. tłum.]: mówiąc w skrócie, finansowanie przedsiębiorstw opiera się, jak się zdaje, na prognozach punktowych, nie zaś na prognozach rozkładów; stąd, jeżeli zadaje się perturbacje prognoz przepływów pieniężnych, powiedzmy w modelu wyceny Gordona, poprzez zmianę ustalonego – i znanego – wzrostu (a także innych parametrów) na zmieniane w sposób ciągły przyrosty (szczególnie w przypadku rozkładów gruboogonowych), to firmy robiące wrażenie „kosztownych” lub charakteryzujące się szybkim wzrostem, ale mające niskie przychody, mogłyby znacznie zwiększyć swoją oczekiwaną wartość, a więc coś, co rynek wycenia heurystycznie, bez żadnych wyraźnych przyczyn.

Wniosek i podsumowanie: Coś, co establishment ekonomiczny zwykł pomijać, to sytuacja, w której mając właściwy model (co jest bardzo daleko idącym założeniem), ale zachowując niepewność co do wartości parametrów, będzie w konsekwencji niezmiennie prowadzić do wzrostu kruchości, przy występowaniu wypukłości i nieliniowości.

Dajmy sobie spokój z małymi prawdopodobieństwami

A teraz istota sprawy – wychodząc poza ekonomię – a mianowicie bardziej ogólny problem z prawdopodobieństwem i jego błędnym pomiarem.

JAK GRUBE OGONY (EKSTREMISTAN) BIORĄ SIĘ Z NIELINIOWYCH REAKCJI NA PARAMETRY MODELU

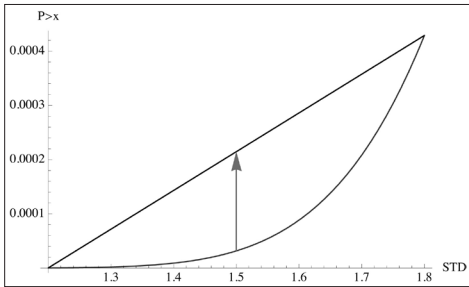
Rzadkie zdarzenia mają pewną właściwość – były pomijane do momentu pisania tej książki. Im więcej niepewności co do wartości parametrów występuje w modelu przeznaczonym do wyznaczania prawdopodobieństw, tym silniejsza tendencja do niedoszacowania małych prawdopodobieństw. Po prostu, małe prawdopodobieństwa są wypukłe względem błędów obliczeniowych, tak jak przelot samolotu jest wypukły względem błędów i zakłóceń (pamiętajmy, że przelot będzie trwał dłużej, nie krócej). Im więcej źródeł zakłóceń zapomni się wziąć pod uwagę, tym dłuższy będzie lot samolotu w stosunku do naiwnego szacunku.

Wszyscy wiemy, że po to, żeby obliczyć prawdopodobieństwo na podstawie klasycznego rozkładu normalnego, potrzebujemy znajomości parametru zwanego odchyleniem standardowym – lub czegoś podobnego, co charakteryzuje rozrzut lub dyspersję wyników. Ale efektem niepewności, odnoszącej się do owego odchylenia standardowego, jest zwiększanie się małych prawdopodobieństw. Dla przykładu, dla odchylenia zwanego „trzy-sigmowym” i zdarzeń, które powinny zrealizować się raz na 740 obserwacji, prawdopodobieństwo rośnie o 60%, jeśli powiększa się odchylenie standardowe o 5%, zaś spada o 40%, jeśli zmniejszamy odchylenie standardowe o 5%. Stąd, jeśli wasz błąd to średnio zaledwie 5%, niedoszacowanie wynikające z zastosowania modelu naiwnego wyniesie około 20%. Olbrzymia asymetria, ale to jeszcze nie wszystko. Sytuacja pogarsza

się, jeśli tropi się odchylenia „sześciosigmowe” (niestety, nagminnie częste w ekonomii): asymetria rośnie pięć razy bardziej. Im rzadsze zdarzenie (tzn. im większa „sigma”), tym gorszy efekt z tytułu niewielkiej tak naprawdę niepewności, co do tego, co umieścić w równaniu. Dla zdarzeń, które charakteryzują się 10 sigmami, różnica rośnie do ponad 10 miliardów. Posłużmy się następującym ciągiem myślowym by wykazać, jak coraz to mniejsze prawdopodobieństwa wymagają coraz większej dokładności obliczeń. Im mniejsze prawdopodobieństwo, w tym większym stopniu małe, nawet bardzo małe zaokrąglenia w obliczeniach czynią asymetrię kompletnie nieistotną statystycznie. Dla małych, naprawdę bardzo małych prawdopodobieństw, niezbędna jest niemal nieskończona dokładność, co do wartości parametrów; najmniejsza tam nawet niepewność powoduje chaos. Są one silnie wypukłe względem perturbacji. Zastosowałem tę argumentację żeby, w pewnym sensie, wykazać, że nie da się obliczyć małych prawdopodobieństw, nawet jeśli dysponuje się właściwym modelem – którym przecież nie dysponujemy.

Analogiczne rozumowanie odnosi się do szacowania prawdopodobieństw metodami nieparametrycznymi, na podstawie dawnych częstości. Jeśli prawdopodobieństwo staje się bliskie 1/wielkość próby, błąd szacunku eksploduje.

Tłumaczy to w oczywisty sposób błąd popełniony w Fukushima. Również w przypadku Fannie Mae. Podsumowując: małe prawdopodobieństwa rosną w przyspieszającym tempie wraz ze wzrostem parametru używanego w obliczeniach.



Rys. 38. Prawdopodobieństwo jest wypukłe względem odchylenia standardowego w modelu Gaussa. Wykres ilustruje wpływ odchylenia standardowego na $P > x$ i porównuje $P > 6$ z odchyleniem

standardowym równym 1,5, z $P > 6$, dla którego zakładamy, że odchylenie standardowe jest kombinacją liniową 1,2 i 1,8 (w tym przypadku $a(1) = 1/5$).

Niepokojące jest to, że perturbacje σ wpływają na ogon rozkładu w sposób wypukły; ryzyka portfela inwestycyjnego, który jest wrażliwy na ogony rozkładu, wzrosną niepomrotnie. Tak więc, nadal jesteśmy w świecie Gaussa! Taka niepewność o skutkach wybuchowych nie wynika w naturalny sposób z grubych ogonów rozkładu, a jedynie z niewielkich niedokładności, co do przyszłej wartości parametru. To w istocie problem poznawczy! Toteż ci, którzy używają takich modeli, jednocześnie uznając istnienie niepewności co do wartości parametrów, dopuszczają się w sposób konieczny poważnej niespójności¹.

Oczywiście, niepewność eksploduje nawet bardziej, kiedy odtwarzamy nie-Gaussowskie warunki świata rzeczywistego, poddając perturbacjom wykładniki ogonowe. Nawet w przypadku rozkładu potęgowego, rezultaty są dotkliwe, szczególnie z tytułu zmienności wykładnika ogonowego, ponieważ ma to ogromne konsekwencje. Tak naprawdę, grube ogon oznaczają niemożność wyliczenia rzadkich zdarzeń, i niewiele więcej.

¹ Idąc dalej, pokazuje to wady koncepcji „niepewności w sensie Knighta”, ponieważ skoro *wszystkie ogony rozkładów* podlegają niepewności z tytułu najmniejszych perturbacji a ich efekt jest dotkliwy w świecie grubych ogonów, tak samo jest w życiu gospodarczym.

WZMAGAJĄC NIEPEWNOŚĆ (FUKUSHIMA)

Korzystając z wcześniejszego stwierdzenia, że estymacja pociąga za sobą błąd, rozszerzmy to rozumowanie: błędy obarczone są błędami; te z kolei mają błędy. Biorąc ten efekt pod uwagę, wszystkie małe prawdopodobieństwa rosną niezależnie od stosowanego modelu – nawet w modelu Gaussa – aż do momentu osiągnięcia grubych ogonów i wystąpienia efektów potęgowych (nawet tak zwanej nieskończonej wariancji), kiedy efekty niepewności wyższych rzędów są duże. Będzie tak nawet biorąc standardowy rozkład Gaussa z odchyleniem standardowym σ , które ma proporcjonalny błąd $a(1)$; $a(1)$ ma z kolei błąd rzędu $a(2)$ itd. Błąd zależy więc od stopy zmian błędu, przechodząc od $a(n-1)$ do $a(n)$; jeśli są to stałe proporcje, to wtedy zbiegamy do rozkładu z bardzo grubym ogonem. Jeśli te proporcjonalne błędy zmniejszają się, nadal mamy do czynienia z grubymi ogonami. Jaki by to nie był przypadek, błąd nie jest korzystnym czynnikiem dla małego prawdopodobieństwa.

To smutne, że skłonienie ludzi do zrozumienia, że każda miara obarczona jest błędem, jest niemal niemożliwe – katastrofa w Fukushima, co do której sądzono, że może zdarzyć się raz na milion lat, mogła mieć przypisane prawdopodobieństwo wystąpienia raz na 30 lat, gdyby odpowiednio przefiltrowano różne warstwy niepewności.