

Dr hab. Michał Krawczyk, prof. ucz.
Uniwersytet Warszawski

Recenzja pracy doktorskiej „Wrażliwość na zmiany prawdopodobieństwa w decyzjach ryzykownych. Rola kontekstu i zdolności numerycznych.”

Trzonem przedstawionej do oceny praca są trzy projekty empiryczne dotyczące wyborów w warunkach ryzyka. Ściślej biorąc, Autor modeluje wrażliwość na prawdopodobieństwo wystąpienia wyniku (niekoniecznie pieniężnego). W normatywnym modelu maksymalizacji oczekiwanej użyteczności, prawdopodobieństwa te bezpośrednio traktuje się jako wagi decyzyjne – model jest liniowy względem prawdopodobieństw i zagadnienie wrażliwości na ich zmiany jest puste. Empirycznie jednak, model ten dość łatwo sfalsyfikować: większość badanych wydaje się silniej reagować na zmiany prawdopodobieństwa w pobliżu naturalnych punktów odniesienia: całkowitej pewności i całkowitej niemożliwości wystąpienia danego – dobrego lub złego – wyniku.

Choć zarówno empiryczna, jak i teoretyczna literatura przedmiotu jest bardzo rozwinięta, we wszystkich badaniach raportowanych w pracy są istotne elementy nowatorskie. W szczególności, w pierwszym projekcie Autor obserwował stawki obstawiane w „zakładach prostych”, podczas gdy przytłaczająca większość eksperymentów opiera się na wyborach między loteriami. W drugim i trzecim projekcie analizował wybory związane z efektami ubocznymi leków, podczas gdy niemal cała literatura przedmiotu skupia się na wynikach pieniężnych. Na pochwałę zasługuje i to, że przed przeprowadzeniem badań zostały one wprowadzone do publicznego rejestru, co zwiększa wagę osiągniętych wyników – zbyt często bowiem w badaniach społecznych (a być może i poza nimi) hipotezy tworzone są ad hoc po obejrzeniu wyników. Uwagę zwraca też wyrafinowana analiza zebranych danych, wykorzystująca metody bayesowskie, regresje hierarchiczne (modele mieszane), symulacje Monte Carlo. To wszystko świadczy o podziwu godnej dojrzałości młodego badacza.

Praca nie jest oczywiście wolna od większych i mniejszych słabości. Zacznę od formalnych. Po pierwsze, moim zdaniem, wypada jasno wskazać zaraz na wstępie jak poszczególne rozdziały mają się do publikacji złożonych przez Autora w pismach, do *working papers* itd. Tu dowiadujemy się na stronie 75 (!), że badanie 3 jest częścią artykułu złożonego do *Journal of Behavioral Decision Making*, a i to tylko dlatego, że Autor informuje, że wykonał je na życzenie recenzentów w tymże piśmie. Nie wiadomo dokładnie jaki jest zakres tego artykułu, czy ma współautorów (!), czy inne fragmenty pracy także zostały – w wersji anglojęzycznej i może zmienionej – opublikowane itd. Taka transparentność, która jest standardem dla doktoratów (z ekonomii) w znanych mi dobrych wydziałach na zachodzie Europy, ułatwia recenzentom i czytelnikom (d)ocenienie wartości dodanej wykonanych badań.

Po drugie, stylistyka kuleje. Gdy czytam o „oryginalnie” (a nie: pierwotnie, początkowo, wstępnie) przyjętych założeniach, „ewaluacji” (a nie: wycenie czy ocenie) loterii, „nisko prawdopodobnych” (a nie mało prawdopodobnych) stratach, „marginalnych” (a nie: krańcowych) użytecznościach, „pozytywnych” czy „negatywnych” (a nie: dodatnich czy ujemnych, jakby wynikało z kontekstu) a nawet „korespondujących” z czymś (a nie:

odpowiadających czemuś) wynikach, to zaczynam myśleć, że może byłoby łatwiej dla wszystkich gdyby praca była po angielsku? Wyrażenia w rodzaju „zob. również Gonzalez i Wu (1999) dla formalnej definicji” są już zupełnie nieznośne. Tak niestety brzmią prace studentów, którzy skorzystali z serwisu Tłumacz Google. Czy zamiast „Z uwagi na fakt, że parametr δ kontroluje elewację PWF...” nie można napisać „Ponieważ to, czy PWF przebiega nisko czy wysoko zależy od (wartości) parametru δ ”? Jeśli „uwaga do strat (ang. *loss attention*), tj. przydzielenie stosunkowo większej uwagi stratom relatywnie do zysków, została przejawiona przez około 94% uczestników” – to wychodzi niestety na to, że w pracy „zostało przejawione przydzielenie relatywnie mniejszej uwagi” elementarnej poprawności stylistycznej. W rezultacie niektórych fragmentów przeglądu literatury po prostu nie daje się czytać. Dalej jest lepiej, choć np. wyrażenie „pomiar wewnątrz jednej osoby” (chodzi o *within-subject design*, s. 61) brzmi na mój gust nazbyt dosłownie.

Mam również dwa istotne zastrzeżenia merytoryczne. Pierwsze dotyczy braku zachęt finansowych dla badanych – wynagrodzenie nie zależało bowiem od ich decyzji. Autor wspomina o hipotetyczności jako problemie, ale poświęca mu bodaj mniej niż pół akapitu. Tymczasem moim zdaniem (które zapewne kształtowane jest przez mój ekonomiczny, nie psychologiczny *background*), jest to w przypadku raportowanych badań problem zasadniczy, ponieważ napotykanne w nich przez badanych zadania były długie i monotonne, nie było w nich elementu interakcji czy konkurowania, słowem: badani doprawdy nie mieli powodu się starać. Może zatem niespójność wyborów świadczyła przede wszystkim o nudzie? Może badani obstawiali nisko nie z powodu awersji do ryzyka, a jedynie by oszczędzić sobie wysiłku wpisywania liczb trzycyfrowych? Tego rodzaju – mało ekscytujących, więc niemiłych autorowi – wyjaśnień nie można lekceważyć. Minimalną reakcją byłoby przetestowanie wynikających z nich przewidywań. Przykładowo, jeśli badani obstawiali kwoty zapisywane przy pomocy powtórzeń tej samej cyfry, np. 99 czy 333, to wiele mówiłoby o chęci przeklikania przez eksperyment szybciej, raczej bardzo mało o stosunku wobec ryzyka. Jeszcze lepszą reakcją Autora byłoby zreplikowanie wyników z uwzględnieniem takich przypuszczeń, np. gdyby minimalny zakład wynosił 100, nie 0 punktów, to nawet najleniwszi badań musieliby tak czy siak wpisać trzy cyfry. Zdecydowanie najlepiej jednak byłoby przeprowadzić podobne badania z realnymi konsekwencjami wyborów. W przypadku „zakładów prostych”, to właśnie bym Autorowi serdecznie polecał. Rzecz prosta, dla efektów ubocznych nie byłoby to już takie łatwe.

Drugie zastrzeżenie dotyczy ostatniego badania. Czy przy takiej jego konstrukcji „wysoko afektywne” skutki uboczne to nie są po prostu poważne skutki uboczne? Jeśli więc prowadzą one do innych wyborów niż te „nisko afektywne”, to nie może to być po prostu efekt wielkości stawki? Jeśli się nie mylę, nie ma możliwości rozróżnienia tych dwóch, a priori prawdopodobnych, wyjaśnień.

W toku lektury poczyniłem także liczne drobniejsze uwagi.

s. 7 Trudno zgodzić się z twierdzeniem, że obstawianie wyniku meczu piłki nożnej to wybór w warunkach ryzyka – prawdopodobieństwo poszczególnych wyników zdecydowanie NIE jest dane. Gdyby było, zapewne zakładów byłoby o wiele mniej. Podobnie indywidualne prawdopodobieństwa powikłań po szczepieniu może znacząco (ale nie wiadomo dokładnie o ile) odbiegać od populacyjnego. Te same zastrzeżenia miałbym do twierdzenia jakoby prawdopodobieństwo wygrania meczu przez daną drużynę można było z góry sprawdzić w internecie (s. 8) i kolejnych przykładów na s. 14.

s. 9 wbrew podanemu przykładowi zakłady sportowe zwykle NIE są symetryczne w podanym sensie. Kto postawiłby na Legię Warszawa przeciwko FC Barcelona, gdyby można było wygrać jedynie postawioną kwotę? Chyba nawet najbardziej lojalni (i nierozgarnięci) fani nie byłiby skłonni do zawarcia takiego zakładu.

W rezultacie, zarówno w literaturze statystycznej jak i dotyczącej hazardu zawarcie „*simple bet*” (odpowiednik terminu „zakład prosty”) wcale nie musi oznaczać, że jeśli wygramy, to dokładnie tyle ile postawiliśmy, por. np. definicje w książkach

Harper, W. L., & Hooker, C. A. (Eds.). (1973). *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science: Volume III* (s. 190).

czy

Barboianu, C. (2006). *Probability Guide to Gambling: The Mathematics of Dice, Slots, Roulette, Baccarat, Blackjack, Poker, Lottery and Sport Bets*. INFAROM Publishing (s. 121).

Nadto z niezwykle rzadkimi wyjątkami (np. mecze finałowe mistrzostw świata, Ligi Mistrzów) mecze piłki nożnej mogą zakończyć się także remisem, a przecież „zakład prosty” ma mieć wg Autora tylko dwa możliwe wyniki. Oczywiście ta rozprawa nie traktuje o piłce nożnej, ale ogólnie ten przykład raczej zaciemnia obraz i podważa twierdzenie Autora o powszechności „zakończeń prostych” w przyrodzie (z którego wywodzi zasadność przyjętego planu badawczego).

s. 15 Rozumiem *gender balance*, ale sugestia jakoby Pascal i Fermat myśleli o hazardzistce (raczej niż hazardziście) jest surrealna.

s. 18 Ściśle biorąc 2.4 nie jest równaniem.

s. 18 (wzór 2.6.) Tu występują niezdefiniowane operacje mnożenia przez skalar oraz dodawania. Nie wiadomo czy a , b , c to te same zwierzęta co A , B , C czy też coś całkiem innego. Nie wiadomo czym na wyrzucenie bez śladu zasłużyło sobie równanie 2.5.

s. 18 Nie bardzo rozumiem co ma oznaczać „również był oparty na pojęciu oczekiwanej użyteczności.” – czemu „również”? I to raczej funkcjonal oczekiwanej użyteczności opiera się na aksjomatach, nie odwrotnie.

s. 21 Nie jest jasne, co miałoby dokładnie znaczyć, że coś staje się „coraz bardziej liniowe”, ale doskonały brak wrażliwości na prawdopodobieństwa w jakimś przedziale dałby płaską, a zatem liniową PWF, znakiem tego zdanie „W miarę zbliżania się wartości tych parametrów do jedynki PWF staje się coraz bardziej liniowe, czyli wrażliwość na prawdopodobieństwa rośnie.” jest podejrzan.

s. 34 „Wykazali, że model, który najlepiej opisywał wybory osób badanych, zakłada oddzielne parametry elewacji PWF dla zysków i strat oraz parametr awersji do ryzyka zablokowany na wartości jeden, tj. brak awersji do ryzyka.” – sądzę, że chodziło jednak o awersję do straty, nie ryzyka.

Rys. 4.2, panel D Mylne oznaczenie. W legendzie winny być koła, nie linie.

Równanie 4.2 Ten model jest wysoce podejrzany. Można przypuszczać, że absolutne błędy będą większe dla większych wartości b_i . Ponadto, gdy b_i jest bliskie dolnemu (górnemu)

ograniczeniu, obserwowana wartość może odchyłać się od niego wyłącznie w górę (dół). Czy założenie o multiplikatywnych, nie addytywnych błędach nie miałoby więcej sensu?

s. 44 Chyba nie rozumiem, co miałyby oznaczać zakłady „wewnętrznie sprzeczne”.

s. 52 Autor wydaje się sugerować, że literatura dotycząca wpływu afektu integralnego na wazenie prawdopodobieństw porównywała loterie pieniężne z loteriami dotyczącymi efektów ubocznych stosowania leków. Nie jest to prawdą, vide

Berns, G. S., Capra, C. M., Moore, S., & Noussair, C. (2007). A shocking experiment: New evidence on probability weighting and common ratio violations. *Judgment and Decision Making*, 2(4), 234–242.

czy

Krawczyk, M. W. (2015). Probability weighting in different domains: The role of affect, fungibility, and stakes. *Journal of Economic Psychology*, 51, 1-15.

s. 57 „Po pierwsze, w poprzednich badaniach kontekst niepieniężny był zawsze stosunkowo silnie nacechowany afektywnie i porównywany ze względnie słabo nacechowanym afektywnie kontekstem niepieniężnym” – chyba ostatnie „nie-” jest niecelowe.

s. 64 Wiem co oznacza estymacja metodą największej wiarygodności i na czym polegają testy oparte o iloraz wiarygodności, ale nie bardzo wiem co miałyby oznaczać maksymalizowanie ilorazu wiarygodności i czemu miałyby służyć.

s. 65 „Jednym ze sposobów na przyspieszenie tych obliczeń jest wybór odpowiednich rozkładów a posteriori” – chyba chodziło o rozkłady a priori.

s. 74 Nie jest dla mnie jasne skąd przekonanie, że bardziej negatywne emocje skutkować będą mniej spójnymi wyborami.

s. 85 Tytuł podsekcji „Poziom wykonania w zadaniu zakłady” jest bardzo brzydki, a meritum – podejrzané. Wiadomo z góry, że model jest niepoprawny. Przykładowo, musi występować interakcja między parametrami, w szczególności brak spójności szkodzi znacznie mniej, gdy ktoś i tak wykazuje silną awersję do ryzyka.

s. 120 Trudno powiedzieć czy osoba w śpiączce może w ogóle być zdenerwowana, a tak by wynikało z tego fragmentu, choć widzę, że to kwestia niefortunnego tłumaczenia z niemieckiego.

s.123 Nie zrozumiałem szczegółów procedury konstruowania wyborów.

Wszystkie te mniejsze i większe zastrzeżenia nie podważają metodologicznej i empirycznej wartości pracy. Stwierdzam zatem, że **przedstawiona do oceny rozprawa spełnia wymagania określone w art. 13 Ustawy z dnia 14 marca 2003 (z późniejszymi zmianami) o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki. Wnioskuje o dopuszczenie mgr. Kamila Fuławki do dalszych etapów przewodu doktorskiego.** Mając na uwadze wysoki poziom merytoryczny rozprawy i udowodnioną przez Autora zdolność do publikowania w międzynarodowych pismach, poddaję także Radzie Instytutu Psychologii pod rozwagę możliwość wystąpienia do J.M. Rektora SWPS o wyróżnienie pracy.