Jerzy Marzec

Wydział Zarządzania
Uniwersytet Ekonomiczny w Krakowie

Stanisław Czyż[1]

Physical Activity, Sport and Recreation
North-West University (South Africa)

Piotr Styrkowiec

Instytut Psychologii
Uniwersytet Wrocławski

Model probitowy dla danych grupowych w analizie skuteczności rzutu do kosza[2]

Streszczenie

W niniejszym artykule zaprezentowano zastosowanie modelu probitowego do opisu skuteczności rzutu do kosza. Na podstawie danych z powtarzalnego eksperymentu zweryfikowano wiele hipotez badawczych. Celem badania było określenie m.in. tego, czy uzyskanie wysoce wyspecjalizowanych umiejętności (*especial skills*) zależy od cech zawodnika (wieku, okresu treningu, skuteczność jego rzutów itp.). Zweryfikowano także hipotezę dotyczącą tego, że wskazówki wizualne (*visual cues*) mają wpływ na zdiagnozowanie *especial skills*. Badania pokazały, że okres treningu, osiągnięcia i skuteczność w meczach ligowych mają silny wpływ na prawdopodobieństwo trafienia do kosza. Dane potwierdziły także nieliniowy i jednocześnie negatywny związek między odległością, z której wykonano rzut, a skutecznością rzutu. Używanie okularów, które zakłócają korzystanie z *visual cues*, istotnie zmniejsza skuteczność rzutów z każdej pozycji. Badania nie potwierdziły, że wykonanie rzutu z linii rzutów wolnych (4,57 m) istotnie zwiększa szanse trafienia w stosunku do rzutów wykonanych z innych pozycji. W konsekwencji hipoteza o występowaniu *especial skills* nie została potwierdzona.

[1] Autor przeprowadził główne badania naukowe prezentowane w niniejszym artykule, będąc zatrudnionym w Akademii Wychowania Fizycznego we Wrocławiu (Wydział Wychowania Fizycznego).

[2] Publikacja została dofinansowana/sfinansowana ze środków przyznanych Wydziałowi Zarządzania Uniwersytetu Ekonomicznego w Krakowie w ramach dotacji na utrzymanie potencjału badawczego.

Słowa kluczowe: model probitowy dla danych grupowych, wysoce wyspecjalizowane umiejętności, skuteczność rzutu do kosza

1. Wstęp

Celem niniejszego artykułu jest zaprezentowanie wykorzystania dychotomicznego modelu probitowego w badaniach, w których dane statystyczne pozyskano w ramach powtarzalnego eksperymentu. Przykład empiryczny dotyczy zagadnienia z zakresu sportu i psychologii eksperymentalnej. W ramach powyższego modelu na podstawie tzw. danych grupowych przeprowadzono testowanie hipotez o charakterze empirycznym dotyczących skuteczności rzutów do kosza.

Dychotomiczny model probitowy jest standardową konstrukcją z klasy modeli dyskretnego wyboru (ang. *quantal response*, *discrete choice models*). Intensywny rozwój tych modeli rozpoczął się już w latach 70. ubiegłego wieku[3]. W ostatniej dekadzie pojawiło się wiele monografii, w których kwestie metodyczne są omawiane w kontekście praktycznych i zaawansowanych badań ekonomicznych[4]. W literaturze polskojęzycznej jest zaledwie kilka takich pozycji. Jedną z nich jest nowoczesny podręcznik akademicki *Mikroekonometria: modele i metody analizy danych indywidualnych*[5], w którym w przystępny sposób przedstawiono podstawowe modele mikroekonometrii w kontekście bardzo ciekawych zastosowań zaczerpniętych z najlepszej literatury przedmiotu. Powyższe modele są także stosowane w statystyce sportu. Przegląd ich zastosowań w różnych dyscyplinach sportowych można znaleźć na łamach „Journal of Quantitative Analysis in Sports".

Model dyskretnego wyboru rozważany w niniejszym artykule może mieć zastosowanie – ze względu na swą strukturę – w badaniach ankietowych gospodarstw domowych lub grup konsumentów. W badaniach tych respondenci są charakteryzowani przez pewien ustalony zestaw cech (np. wielkość gospodarstwa, miejsce zamieszkania, płeć

---

[3] G.S. Maddala, *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge 1983; T. Amemiya, *Advanced Econometrics*, Harvard University Press, Cambridge 1985.

[4] Zob. np. J.S. Cramer, *Logit Models From Economics and Other Fields*, Cambridge University Press, Cambridge 2003; A.C. Cameron, P.L. Trivedi, *Microeconometrics*: *Methods and Application*, Cambridge University Press, New York 2005; D.A. Hensher, J.M. Rose, W.H. Greene, *Applied choice analysis: A primer*, Cambridge University Press, Cambridge 2005; R. Winkelmann, S. Boes, *Analysis of Microdata*, Springer-Verlag, Heidelberg 2006.

[5] *Mikroekonometria: modele i metody analizy danych indywidualnych*, red. M. Gruszczyński, Oficyna a Wolters Kluwer business, Warszawa 2010.

głowy rodziny, wielkość dochodów mierzoną na skali porządkowej). Częste powtarzanie się wartości tych cech umożliwia pogrupowanie respondentów i łączną analizę częstości odpowiedzi na postawione w ankiecie pytania. Niniejszy przykład, dotyczący skuteczności rzutu do kosza, pozwoli spojrzeć na szerokie możliwości modelowania ekonometrycznego w kontekście stawianych hipotez empirycznych.

2. Rzut do kosza – opis problemu

Zastanówmy się nad tym, w jaki sposób człowiek nabywa umiejętność nieprzypadkowego trafienia piłką do kosza zawieszonego na wysokości 3,05 m. Jest to możliwe dzięki *especial skills*, które można przetłumaczyć jako „umiejętności wyjątkowe w swojej klasie” czy „wysoce wyspecjalizowane umiejętności”. Jako pierwsi pojęcie to wprowadzili do literatury K.M. Keetch i inni[6]. Specjalizacja taka dokonuje się przez tysiące (dziesiątki lub setki tysięcy) powtórzeń danej umiejętności. Powtórki są wykonywane w tych samych warunkach zewnętrznych, co oznacza, że ćwiczący używa tych samych przyrządów (o tej samej masie, wymiarach itp.), ćwiczy na identycznych matach, bieżniach, rzuca, skacze, uderza przedmiotami o identycznych wymiarach lub masie do celu umieszczonego na tej samej wysokości, w tej samej odległości od ćwiczącego itd. W literaturze przedmiotu sugeruje się, że zjawisko *especial skills* może się w pełni ujawnić w przypadku rzutów z pozycji „osobiste” (4,57 m), które na treningach zawodnicy wykonują w niezliczonej liczbie.

W celu pomiaru zjawiska *especial skills* przeprowadzono eksperyment. Polegał on na pomiarze skuteczności rzutu koszykarza w zależności od dwóch czynników – pozycji mierzonej odległością od obręczy kosza i posiadania specjalnych okularów osłabiających widzenie. Wprowadzenie okularów miało na celu sprawdzenie tego, w jakim stopniu wskazówki wizualne (*visual cues*) pomagają skutecznie skierować piłkę do celu. Uczestnikami eksperymentu byli młodzi zawodnicy z drugoligowego klubu WKS Śląsk. Eksperyment wykonano w czterech turach. Każdy z zawodników oddał łącznie 1050 rzutów, które wykonał z siedmiu odległości. Połowa rzutów została wykonana, gdy zawodnik miał założone okulary o parametrach +4 dioptrie. Okulary te w przypadku zawodników nieposiadających wad wzroku zniekształcają pole widzenia. Każdy z dziesięciu zawodników wykonał 75 powtórzeń tego doświadczenia dla każdej

[6] K.M. Keetch, R.A. Schmidt, T.D. Lee, D.E. Young, *Especial Skills: Their Emergence with Massive Amounts of Practice*, „Journal of Experimental Psychology: Human Perception and Performance” 2005, vol. 31, s. 970–978.

z 14 kombinacji dwóch czynników, tj. pozycji rzutu i okularów. Szczegółowy opis problemu *especial skills* i eksperymentu zaprezentowano w opracowaniu S.H. Czyża, J. Marca i P. Styrkowca[7].

3. Uwagi na temat modelu dla danych grupowych

W ujęciu statystycznym przedmiotem obserwacji jest zmienna o rozkładzie dwupunktowym. Obserwujemy trafienie do kosza albo „pudło". Jej zaobserwowana wartość jest oznaczona przez y_{ijgl}, gdzie i to numer eksperymentu (doświadczenia) ($i = 1,\ldots,75$), $j = 1$ albo 0, co oznacza odpowiednio wykonanie rzutu w specjalnych zniekształcających obraz okularach albo bez ich użycia, g określa pozycję rzutu ($g = 1,\ldots,7$), a l to numer zawodnika ($l = 1,\ldots,10$). Łączna liczba obserwacji zmiennej y_{ijgl} wynosi zatem 10 500. Biorący udział w eksperymencie koszykarze są charakteryzowani przez zestaw indywidualnych cech (np. okres trwania treningu koszykówki, liczbę rozegranych meczów w pierwszym składzie drużyny itp.). Oczywistą kwestią jest fakt, że cechy te są niezmienne w każdym doświadczeniu (rzucie). W konsekwencji rozważany zbiór obserwacji ma charakter tzw. danych grupowych. Według C. Gourieroux[8], z danymi grupowymi (ang. *group data*) mamy do czynienia w sytuacji, gdy w próbie dla różnych wartości zmiennej endogenicznej obserwujemy identyczne kombinacje wartości zmiennych egzogenicznych. Dane takie uzyskuje się przez agregację pojedynczych obserwacji – danych indywidualnych. M. Gruszczyński[9] nazywa je indywidualnymi makrodanymi, a model dychotomiczny dla tych danych – modelem frakcji. W omawianym przypadku obiektem analizy statystycznej będzie efekt rzutu – trafił albo nie trafił – wykonanego przez danego zawodnika z ustalonej pozycji i w okularach albo bez nich. W konsekwencji model statystyczny jest oparty na obserwacjach y_{it} uzyskanych dla $T = 140$ obiektów w wyniku przeprowadzonych eksperymentów w liczbie $N = 75$. Wartość $y_{it} = 1$ oznacza sukces, czyli trafienie do kosza przez wybranego zawodnika

[7] S.H. Czyż, J. Marzec, P. Styrkowiec, *Free throw shot proficiency in basketball – a probability model*, Working Papers, 2012, https://e-uczelnia.uek.krakow.pl/pluginfile.php/68421/mod_page/content/12/Artykuly/2012_Free_throw_shot_proficiency_in_basketball_-_a_probability_model_CzyzMarzecStyrkowiec_ver2012_09.pdf (odczyt: 10.04.2015).

[8] C. Gourieroux, *Econometrics of Qualitative Dependent Variables*, Cambridge University Press, Cambridge 2000.

[9] M. Gruszczyński, *Model zmiennych jakościowych dwumianowych*, w: *Mikroekonometria: modele i metody*…, op.cit., s. 53–102.

w sytuacji określonej przez odległość do kosza i posiadanie albo nieposiadanie okularów, $y_{it} = 0$ – porażkę dla $t = 1,\dots,T$.

W literaturze przedmiotu model dychotomiczny ma strukturę hierarchiczną z regresją dla ciągłej zmiennej ukrytej. W niniejszych badaniach zmienna ta informuje o zdolnościach (umiejętnościach) trafienia do celu przez poszczególnych zawodników. Jeżeli zdolności przewyższają umowną wartość progową, to obserwuje się trafienie do kosza, w przeciwnym przypadku rzut kończy się porażką. Można przyjąć, że poszczególne rzuty są realizacjami niezależnych zmiennych losowych o rozkładach dwupunktowych. Zmiana pozycji rzutu lub nieposiadanie okularów zmienia parametr tego rozkładu. Ponadto, cechy indywidualne zawodników wpływają na rozkład tej zmiennej losowej.

W przypadku danych grupowych model dychotomiczny ma następującą postać[10]

$$y_{it} = I_{[0;+\infty)}\left(z_{it}\right) \qquad \text{gdzie} \quad z_{it} = x_t \cdot \beta + \varepsilon_{it}, \tag{1}$$

gdzie $I_\Omega\left(\omega\right)$ jest funkcją charakterystyczną zbioru, tzn. jeżeli $\omega \in \Omega$, to $I_\Omega\left(\omega\right) = 1$, 0 – w przeciwnym przypadku. Ponadto, β to wektor nieznanych parametrów, wiersz x_t jest wektorem-wierszem, który zawiera charakterystyki zawodnika oraz informacje o użyciu okularów i odległości, z której wykonano rzut (i dodatkową zmienną „1"). Założenia dotyczące ε_{it} są standardowe. Dla rozkładów symetrycznych prawdopodobieństwo sukcesu jest równe $p_t = F\left(x_t \cdot \beta\right)$, gdzie $F\ (a)$ jest dystrybuantą w punkcie a jednego z rozkładów prawdopodobieństwa przyjętego dla ε_{it}. W tych badaniach przyjęto, że ε_{it} jest standaryzowaną zmienną o rozkładzie normalnym.

Standardową metodą estymacji parametrów modelu (1) jest metoda największej wiarygodności (MNW). Logarytm funkcji wiarygodności w przypadku T obiektów i N_t doświadczeń ma postać:

$$\ln L\left(\beta; y\right) = \sum_{t=1}^{T} N_t \bar{p}_t \ln\left(F\left(x_t\beta\right)\right) + N_t\left(1 - \bar{p}_t\right)\ln\left(1 - F\left(x_t\beta\right)\right), \tag{2}$$

gdzie $\bar{p}_t = \sum_{t=1}^{N_t} y_{it} \Big/ N_t$ to częstość zaobserwowania sukcesu dla obiektu t w N_t doświadczeniach. Macierz drugich pochodnych cząstkowych w przypadku modeli logitowego i probitowego jest macierzą ujemnie określoną dla dowolnego β[11]. Logarytm funkcji wiarygodności jest zatem funkcją globalnie wklęsłą, więc istnieje jej maksimum i można jednoznacznie wyznaczyć ocenę estymatora dla danej próby[12]. Macierz kowariancji

[10] Zob. np. T. Amemiya, op.cit.; C. Gourieroux, op.cit.

[11] Zob. T. Amemiya, op.cit.

[12] Z wyłączeniem pewnego szczególnego przypadku, tj. separowalności próby; zob. J. Albert, J.A. Anderson, *On the Existence of Maximum Likelihood Estimates in Logistic Regression*, „Biometrika" 1984, vol. 71, s. 1–10.

estymatora, którą uzyskuje się analitycznie, korzystając z odwrotności macierzy informacyjnej Fishera, ma postać:

$$V\left(\hat{\beta}_{MNW}\right)=\left[\sum_{t=1}^{T}\frac{N_t\cdot f\left(x_t\beta\right)^2}{F\left(x_t\beta\right)\cdot\left(1-F\left(x_t\beta\right)\right)}x_t'x_t\right]^{-1}, \tag{3}$$

gdzie $f\left(x_t\beta\right)$ to funkcja gęstości prawdopodobieństwa zmiennej ε_{it}. Jeżeli eksperyment został przeprowadzony tylko raz, tj. $N_t = 1$ dla każdego t, to otrzymujemy model dychotomiczny dla danych indywidualnych. Zauważmy, że posiadanie danych grupowych ($N_t > 1$) powoduje, iż oceny parametrów charakteryzują się większą precyzją niż w drugim przypadku.

Warto wspomnieć o tym, że istnieją alternatywne metody estymacji modelu (1), jednakże metoda MNW wyróżnia się na ich tle, gdyż jest procedurą jednostopniową. Inną metodą estymacji jest uogólniona metoda najmniejszych kwadratów (UMNK) dla aproksymacji modelu (1) w formie równania regresji ze zmienną objaśnianą, która przyjmuje wartości równe funkcji odwrotnej dla $F\left(x_t\beta\right)$, a składnik losowy charakteryzuje się heteroscedastycznością[13]. W najbardziej znanym, podręcznikowym przypadku, tj. regresji logistycznej, zmienna objaśniana jest wyrażona za pomocą logarytmu ilorazu szans, czyli jako $ln\left(\bar{p}_t/\left(1-\bar{p}_t\right)\right)$. Generalnie estymację przeprowadza się dwuetapowo, ale procedura jednostopniowa jest także możliwa, gdy zastosuje się nieliniową ważoną MNK. Wspomniane estymatory UMNK i MNW mają te same własności asymptotyczne.

4. Konstrukcja modelu opisującego rzut do kosza

W zastosowanym modelu probitowym zmienną objaśnianą jest procentowa trafność rzutów do kosza danego zawodnika uzyskana z ustalonych pozycji i w okularach albo bez nich. Zróżnicowanie skuteczności rzutów oczywiście może zależeć od odległości (w_{t1}), z której został wykonany rzut. Wyróżniono następujące odległości: 2,74 m, 3,35 m, 3,96 m, 4,57 m, 5,18 m, 5,79 m oraz 6,4 m. Ponadto, posiadanie specjalnych okularów ($w_{t2} = 1$, gdy posiada) może obniżyć skuteczność, gdyż ich użytkowanie zakłóca postrzeganie odległości między koszem a miejscem rzutu, czyli korzystanie ze wskazówek wizualnych. Jednakże zawodnik zna odległość, z której rzuca do kosza, więc okulary nie powinny być czynnikiem zakłócającym pod warunkiem, że wszystkie składowe procesu psychofizycznego, który skutkuje wykonaniem rzutu, są automatyczne, a więc

[13] Zob. C. Gourieroux, op.cit.; M. Gruszczyński, op.cit.

perfekcyjnie wyuczone. W celu pełnego opisu badanego zjawiska przyjęto dodatkowo, o czym była mowa wcześniej, że zawodnika opisują następujące cechy: okres trwania treningu koszykówki (w latach; w_3), liczba rozegranych meczów w pierwszym składzie drużyny w ostatnim sezonie (w_4), średni czas (w minutach) pobytu na parkiecie w meczach ligowych (w_5), średnia liczba punktów zdobytych w tych meczach (w_6). Nabyte przez lata treningu i współzawodnictwa sportowego umiejętności zawodnika zostały uwzględnione przez: skuteczność rzutów za 2 i 3 punkty (w_7, w_8) oraz trafność rzutów z linii rzutów wolnych (potocznie zwanych rzutami osobistymi; w_9), uzyskaną przez każdego zawodnika w meczach ligowych w ostatnim sezonie. Hipotezy badawcze przyjęły formę:

1) wraz ze zwiększaniem odległości nieproporcjonalnie obniża się zdolność trafienia do kosza (H1),
2) jeżeli zawodnicy używają okularów, które zakłócają korzystanie z *visual cues*, to skuteczność rzutów jest mniejsza niż w przypadku braku okularów (H2),
3) istnieją czynniki charakteryzujące samego zawodnika, które istotnie wpływają na skuteczność rzutu, tj. lata treningu, dotychczasowe osiągnięcia w lidze, skuteczność gry w meczach (H3),
4) bez względu na fakt założenia okularów wykonanie rzutu z pozycji „osobiste" (4,57 m) zwiększa szanse trafienia w stosunku do rzutów wykonanych z innych pozycji, nawet tych z bliższych odległości (H4),
5) istnieje pewna odległość (nieznana *ex ante*), z której szansa trafienia jest największa, i nie jest to koniecznie punkt znajdujący się najbliżej obręczy kosza, tj. w odległości 2,74 m, albo pozycja „osobiste" (H5),
6) jeżeli prawdziwa jest hipoteza o parametryzacji[14], tłumacząca powstawanie *especiall skills*, to w przypadku zawodników, którzy używają okularów zakłócających korzystanie ze wskazówek wizualnych, w mniejszym stopniu pogorszy się skuteczność rzutów z odległości 4,57 m niż skuteczność rzutów z pozostałych odległości (H6).

W celu weryfikacji hipotez skonstruowano odpowiednią postać regresji dla zmiennej ukrytej z_{it}:

$$z_{it} = \beta_0 + \sum_{j=1}^{9} \beta_j w_{tj} + \beta_{10} I_{(4,57)}\left(w_{t1}\right) + \beta_{11}\left(w_{t1}\right)^2 + \beta_{12} w_{t2} I_{(4,57)}\left(w_{t1}\right) + \varepsilon_{it}, \tag{4}$$

gdzie $I_\Omega(\omega)$ jest identyczne z tym we wzorze (1).

[14] K.M. Keetch, T.D. Lee, R.A. Schmidt, *Especial Skills: Specificity Embedded Within Generality*, „Journal of Sport & Exercise Psychology" 2008, vol. 30(6), s. 723–736; G. Breslin, N.J. Hodges, R. Kennedy, M. Hanlon, A.M. Williams, *An especial skill: Support for a learned parameters hypothesis*, „Acta Psychologica" 2010, vol. 134, s. 55–60; G. Breslin, R.A. Schmidt, T. Lee, *Especial Skills: Generality and Specificity in Motor Learning*, w: *Skill Acquisition in Sport, Research Theory and Practice*, red. A.M. Williams, N. Hodges, Taylor & Francis, London 2012.

W tabeli 1 zaprezentowano zestawienie modeli w kontekście postawionych hipotez. Formuła (4) wskazuje, że przyjęto, iż umiejętność (zdolność) skutecznego rzutu jest nieliniową funkcją odległości rzutu (w_{t1}) i posiadania okularów (w_{t2}). Ujemny parametr β_1 i restrykcja $\beta_{10} = \beta_{11} = \beta_{12} = 0$ uprawdopodabniają hipotezę H1, że wraz ze zwiększaniem odległości maleją szanse trafienia do kosza. Ujemny znak β_2 jest warunkiem koniecznym prawdziwości hipotezy H2. Natomiast dodatni parametr β_{10} odzwierciedla przekonanie, że wykonanie rzutu osobistego zwiększa szanse trafienia w stosunku do rzutów wykonanych z innych pozycji (hipoteza H4). W hipotezie H5 decydującą rolę odgrywa parametr β_{11}. Jego dodatni znak oznacza, że istnieje wartość zmiennej w_{t1}, która minimalizuje p_t, zaś ujemny – maksymalizuje. Hipotezę H5 potwierdza ujemna wartość oceny dla β_{11}. Zauważmy, że hipotezy H1 i H4 są szczególnymi przypadkami H5. Wobec tego interesującą kwestią jest sprawdzenie, który z tych efektów jest dominujący. Hipoteza H6 jest koniunkcją hipotez H2 i H4. Prawdziwość hipotezy H6 sprzyja obu hipotezom, a szczególnie H4. Dodatni znak parametru β_{12} jest warunkiem koniecznym i wystarczającym zasadności hipotezy H6. Parametr ten informuje o wzroście skuteczności rzutu z 4,57 m w porównaniu z rzutami z innych odległości pod warunkiem, że w eksperymencie wykorzystano okulary.

Tabela 1. Opis modeli w kontekście rozważanych hipotez

Model	Interpretacja modelu	Restrykcje względem M0	Liczba parametrów
M0	–	–	13
M1	brak efektu „nieliniowej zależności”	$\beta_{10} = \beta_{11} = \beta_{12} = 0$	10
M2	brak efektu „okulary”	$\beta_2 = 0$	12
M3	brak efektu „cechy zawodnika”	$\beta_3 = \ldots = \beta_9 = 0$	6
M4	brak efektu „rzut osobisty”	$\beta_{10} = 0$	12
M5	brak efektu „optymalna odległość”	$\beta_{11} = 0$	12
M6	brak efektu „rzut osobisty w okularach”	$\beta_{12} = 0$	12

Źródło: opracowanie własne.

Przy budowie modeli wykorzystano podejście „od ogółu do szczegółu” (ang. *general to specific modelling*). W celu weryfikacji powyższych hipotez empirycznych, H1–H6, poddano estymacji parametry modelu głównego i sześciu pomocniczych. Model ogólny (M0) wyraża najmniej restrykcyjne (najpełniejsze) przekonanie badacza dotyczące potencjalnych zależności między zmienną objaśnianą a zmiennymi objaśniającymi. Redukcję modelu M0 przeprowadzono z wykorzystaniem testu ilorazu wiarygodności (LR). Dodatkowo, w celu porównania mocy wyjaśniającej rozważanych modeli, które są konstrukcjami niezagnieżdżonymi, tj. M2, M4 i M5 oraz M3 i M5, zastosowano kryterium informacyjne Akaikego.

Podejście to różni się zasadniczo od tego prezentowanego w literaturze poświęconej *especial skills*. W artykułach współautorstwa K.M. Keetch[15], które stanowią punkt wyjścia badań podejmowanych przez innych, zastosowano bardzo prostą procedurę. Posiadając dane identyczne z tymi prezentowanymi w niniejszym artykule, wykorzystano model regresji liniowej do opisu związku między częstością trafienia do kosza a odległością. Następnie, pomijając w próbie informacje o odległości 4,57 m, dokonano osobnej estymacji tegoż modelu dla każdego z badanych siedmiu zawodników w dwóch przypadkach – rzutów wykonanych w okularach albo bez nich. Na ich podstawie wyznaczono prognozy prawdopodobieństw trafienia dla rzutu wolnego (4,57 m) dla każdego z zawodników, które potraktowano jako realizację zmiennej losowej o rozkładzie normalnym i parametrach równych średnim z próby. W dwóch wariantach – wyniki dla rzutów w okularach i bez nich – zastosowano jednostronny test *t*-Studenta dla dwóch średnich, porównując średnią wartość otrzymanych prognoz (po wszystkich zawodnikach) z empiryczną częstością trafienia z odległości 4,57 m. Za ocenę parametru rozproszenia zmiennej losowej przyjęto próbkową wariancję z ośmiu prognoz zmiennej objaśnianej. Potwierdzeniem hipotezy o istotnej roli *especial skills* było odrzucenie hipotezy o równości średnich w przypadku rzutu w okularach przy jednoczesnej przeciwnej konkluzji, gdy rzut wykonano bez okularów. Przytoczona procedura jest bardzo prosta, ale w kontekście niniejszych badań szczególnie istotną kwestią jest to, że nie dopuszcza odstępstwa o liniowości wspomnianej zależności.

5. Wyniki estymacji i testowania hipotez

Syntetyczny opis próby statystycznej uzyskany w wyniku przeprowadzonego eksperymentu zawarto w tabeli 2. Każdy z zawodników, grając w meczach ligowych, wykonywał rzuty ze wszystkich pozycji. Ponadto, koszykarze są zróżnicowani ze względu na siedem wyróżnionych cech.

Szczegółowe informacje dotyczące wyników estymacji najlepszych wariantów modelu (1) zaprezentowano w tabeli 3. Warto zauważyć, że oceny parametrów i błędy estymacji są stabilne pomiędzy poszczególnymi modelami. W następnych tabelach (4 i 5) w formie uzupełnienia pokazano wyniki testu LR oraz ranking modeli według kryterium Akaikego. Wyniki te niosą zgodną informację o mocy wyjaśniającej poszczególnych modeli.

[15] K.M. Keetch, R.A. Schmidt, T.D. Lee, D.E. Young, op.cit.; K.M. Keetch, T.D. Lee, R.A. Schmidt, op.cit.

Tabela 2. Sumaryczny opis zawodników uczestniczących w eksperymencie

i	Lata treningu (w_{t3})	Liczba meczów w lidze (w_{t4})	Średni czas gry w meczu (minuty) (w_{t5})	Średnia liczba punktów w meczu (w_{t6})	Procentowa skuteczność w meczu		
					za 2 pkt (w_{t7})	za 3 pkt (w_{t8})	z linii 4,57 (w_{t9})
1	6	8	17	4,4	42	27	50
2	4,5	16	10	2,3	44	28	60
3	11	11	6	1,3	40	25	57
4	7	24	12	1,9	50	23	44
5	8	23	25	13,6	61	30	63
6	5	14	9	2,9	58	0	63
7	10	23	21	9,7	48	27	67
8	7	15	7	0,9	43	9	56
9	9	16	8	2,8	53	30	67
10	8	24	26	6,1	47	15	71

Źródło: opracowanie własne.

Tabela 3. Wyniki estymacji modeli probitowych

Zmienna	Parametr	Model M0	Model M4	Model M6
		ocena (błąd)	ocena (błąd)	ocena (błąd)
„1”	β_0	0,277* (0,169)	0,217* (0,165)	0,280* (0,169)
Odległość rzutu w metrach (w_{t1})	β_1	0,296 (0,057)	0,328 (0,054)	0,296 (0,057)
Czy ma okulary? (w_{t2})	β_2	–0,273 (0,015)	–0,279 (0,015)	–0,280 (0,014)
Lata treningu (w_{t3})	β_3	–0,039 (0,005)	–0,039 (0,005)	–0,039 (0,005)
Liczba meczów w lidze (w_{t4})	β_4	0,014 (0,002)	0,014 (0,002)	0,014 (0,002)
Średni czas gry (w_{t5})	β_5	–0,008 (0,003)	–0,008 (0,003)	–0,008 (0,003)
Średnia liczba punktów w meczu (w_{t6})	β_6	0,013 (0,006)	0,013 (0,006)	0,013 (0,005)
Skuteczność w meczu za 2 pkt (w_{t7})	β_7	–0,019 (0,002)	–0,019 (0,002)	–0,019 (0,002)
Skuteczność w meczu za 3 pkt (w_{t8})	β_8	0,006 (0,001)	0,006 (0,001)	0,006 (0,001)
Skuteczność w meczu z 4,57 (w_{t9})	β_9	0,018 (0,001)	0,018 (0,001)	0,018 (0,001)
Rzut z linii 4,57 (tak, nie)	β_{10}	0,050* (0,031)	–	0,025* (0,023)
(Odległość)^2	β_{11}	–0,062 (0,006)	–0,066 (0,006)	–0,062 (0,006)
Rzut w okularach z 4,57 (tak, nie)	β_{12}	–0,046* (0,041)	–0,003* (0,030)	–
Ln L (β; y)	–	–6509,62	–6510,01	–6509,82

Symbole „ – ” i „*” oznaczają odpowiednio pominięcie zmiennej w modelu i nieistotność na poziomie 0,05.
Źródło: opracowanie własne.

Hipoteza pierwsza (H1) zakłada, że występuje nieliniowa zależność między zdolnością rzutu do kosza (traktowana jako zmienna ukryta) a odległością, z której został on wykonany. Dane nie potwierdzają, że $\beta_{10} = \beta_{11} = \beta_{12} = 0$, czyli model M1, w którym neguje się prawdziwość H1, jest nieuzasadniony wobec M0. Statystyka LR dla obu modeli wynosi 44,08 przy wartości krytycznej 7,81 i granicznym poziomie istotności (*p-value*) równym około 10^{-8}. Kryterium Akaikego także preferuje model M0 w stosunku do modelu M1. Ten ostatni jest dopiero piąty w tym zbudowanym z siedmiu pozycji rankingu, gdyż zakłada m.in. nieistotność oceny parametru β_{11}, czego zdecydowanie nie potwierdzają dane. W konsekwencji jednakże hipoteza H1 przegrywa z hipotezą H5, czyli efekt istnienia „optymalnej odległości" jest mocniejszy od ogólnego efektu „nieliniowej zależności".

Hipoteza H2, głosząca, że posiadanie okularów pogarsza skuteczność rzutów, została jednoznacznie potwierdzona we wszystkich modelach. Ocena parametru β_2 jest ujemna, charakteryzuje się stabilnością i precyzją szacunku. Usunięcie tej zmiennej z modelu wyjściowego M0 powoduje, że otrzymana w ten sposób specyfikacja M2 znajduje się na przedostatnim miejscu w rankingu.

W odniesieniu do hipotezy H3 zauważmy, że wyniki testowania mocno ją wspierają. Wynik testu LR dla łącznej istotności ocen dla β_3, β_4, β_5, β_6, β_7, β_8 i β_9 zdecydowanie wskazuje na zasadność tej hipotezy. W konsekwencji model M3, w którym przyjęto, że łączny wpływ wspomnianych zmiennych jest statystycznie nieistotny, znajduje się na ostatnim miejscu w rankingu. Ponadto, w modelu ogólnym M0 wartości statystyk *t*-Studenta dla ocen pojedynczych wspomnianych parametrów są większe od 2. Świadczą one o istotności wpływu pojedynczych zmiennych charakteryzujących koszykarza na skuteczność rzutu. W świetle powyższych argumentów hipoteza ta została w pełni potwierdzona. Oceny trzech spośród siedmiu parametrów mają statystycznie ujemny znak. Dotyczy to lat treningu, średniego czasu gry i skuteczności rzutu za 2 pkt w meczach ligowych. W tym kontekście warto zauważyć, że w eksperymencie brali udział młodzi zawodnicy, z doświadczeniem na poziomie klubu drugoligowego. Ponadto, czas ich gry w meczach może zależeć od dyspozycji pozostałych zawodników oraz indywidualnych decyzji trenera. W przypadku ostatniej zmiennej ujemny znak parametru jest dopuszczalny, gdyż podczas gry skuteczność rzutu za 2 pkt zależy silnie od wielu niezależnych okoliczności o charakterze losowym (ustawienia zawodników drużyny przeciwnej itp.).

Hipoteza czwarta, o szczególnej roli pozycji rzutu określonej w punkcie oddalonym od obręczy o 4,57 m, nie została w tych badaniach zdecydowanie potwierdzona. W omawianych modelach oceny parametru β_{10} są, co prawda, dodatnie, ale statystycznie nieistotne w modelach z pierwszego i trzeciego miejsca w rankingu, tj. M6 i M0. Redukcja modelu M0 do modelu M4 jest w pełni zasadna. Wysoka wartość *p-value* ($\approx$0,38) dla testu LR wskazuje na brak wystarczających dowodów na istnienie efektu „rzutu osobistego". Efekt ten uwidacznia się dopiero, gdy w analizie pominie się dowody świadczące o efekcie „nieliniowej zależności" (model M2) lub o efekcie „optymalnej odległości" (M5). Niestety modele M2 i M5, wykluczające oba efekty, są mało prawdopodobne, o czym świadczą m.in. ich odległe miejsca w rankingu, odpowiednio 6 i 4.

Tabela 4. Wyniki testu ilorazu wiarygodności

Hipoteza	Statystyka LR	Liczba restrykcji	Wartość krytyczna testu χ^2 ($\alpha = 0{,}05$)	*P-value*
M1 kontra M0	44,08	3	7,81	$<10^{-8}$
M2 kontra M0	97,41	1	3,84	$<10^{-22}$
M3 kontra M0	134,10	6	12,59	$<10^{-25}$
M4 kontra M0	0,78	1	3,84	0,377
M5 kontra M0	30,42	1	3,84	$<10^{-7}$
M6 kontra M0	0,40	1	3,84	0,525

Źródło: opracowanie własne.

Tabela 5. Ranking modeli według kryterium Akaikego

Model	Interpretacja modelu	Liczba parametrów	AIC	Ranking
M0	–	13	13 045,2	3
M1	brak „nieliniowej zależności"	10	13 083,3	5
M2	brak efektu „okulary"	12	13 140,7	6
M3	brak efektu „cechy zawodnika"	7	13 167,4	7
M4	brak efektu „rzut osobisty"	12	13 044,0	2
M5	brak efektu „optymalna odległość"	12	13 073,7	4
M6	brak efektu „rzut osobisty w okularach"	12	13 043,7	1

Źródło: opracowanie własne.

W hipotezie H5 przyjmuje się, że w kontekście skuteczności istnieje pewna optymalna odległość rzutu. Zauważmy, że choć hipoteza H1 została potwierdzona, to jednak H4 (jako szczególny przypadek H1) nie znalazła silnego poparcia w danych. Ocena parametru β_{11} jest statystycznie ujemna we wszystkich modelach. Próba pominięcia efektu „optymalnej odległości", uwidoczniona w wynikach estymacji modelu M5, nie znajduje uzasadnienia w danych. Test LR dla modeli M5 kontra M0 silnie wskazuje na rzecz tego drugiego. Dane w pełni zatem potwierdzają hipotezę H5. W konsekwencji można wyznaczyć optymalną odległość rzutu, która maksymalizuje prawdopodobieństwo trafienia. W przypadku trzech najlepszych modeli – M6, M4 i M0 – wynosi ona odpowiednio 2,37 m, 2,49 m i 2,37 m, więc jest krótsza od odległości 2,74 m, która stanowiła pierwszy punkt, z którego koszykarze oddawali rzut w niniejszym eksperymencie. Zależność między zdolnością trafienia do kosza (z_{it}) a odległością (w_{t1}) jest zatem nieliniowa, ale jednocześnie ujemnie monotoniczna w zakresie siedmiu rozważanych odległości. Wykryta zależność skutkuje tym, iż wraz ze zwiększaniem w_{t1} prawdopodobieństwo sukcesu szybko maleje do zera. Jednakże w literaturze przedmiotu zwraca się uwagę na fakt, że w przypadku najskuteczniejszych koszykarzy ligi NBA dane pochodzące z meczów (czyli z gry) wskazują na *s*-kształtną zależność między z_{it} a odległością[16]. Kwestia ta jawi się jako ciekawy temat dalszych, pogłębionych badań, które jednakże wymagają posiadania liczniejszej próby ze względu na odległości, z której są wykonywane rzuty.

Weryfikacja hipotezy szóstej opiera się na zbadaniu tego, czy $\beta_{12} > 0$. We wszystkich pięciu modelach uwzględniających efekt „rzutu osobistego w okularach" ocena tego parametru jest ujemna, aczkolwiek często charakteryzuje się słabą precyzją estymacji. Model, w którym pominięto w_{t12}, jest najlepszy według przyjętego kryterium informacyjnego. Test LR dla hipotezy M6 kontra M0 wskazuje na zasadność usunięcia tej zmiennej. Wyniki te zdecydowanie nie potwierdziły prawdziwości hipotezy szóstej.

W nawiązaniu do powyższych negatywnych dowodów dotyczących hipotezy H6 przeprowadzono dalsze pogłębione badania w tym zakresie. Okazało się, że postulowany w tej hipotezie efekt „rzutu osobistego w okularach" zostanie potwierdzony wyłącznie wówczas, gdy w procedurze weryfikacji pominie się jednocześnie dwa efekty – „optymalnej odległości" (H5) i „rzutu osobistego" (H4). Niestety wskutek pominięcia obu tych ważnych zmiennych objaśniających ocena parametru β_1 będzie ujemna, a więc całkowicie niezgodna z intuicją. Reasumując, należy stwierdzić, że w świetle wcześniejszych wyników takie postępowanie byłoby sprzeczne z informacją zawartą w próbie.

6. Podsumowanie

Powyższe rezultaty nie potwierdzają występowania efektu związanego z istotnym zwiększeniem skuteczności rzutu z linii rzutów „osobistych". Także różnica w skuteczności rzutu z tej pozycji, wywołana użyciem okularów zakłócających pole widzenia, nie jest istotnie mniejsza w stosunku do pozostałych odległości. Dominującym efektem jest nieliniowa i silnie ujemna zależność między skutecznością rzutu a odległością. Aspekt ten był *a priori* pomijany w badaniach przedstawianych w literaturze na rzecz hipotezy

[16] Zob. J.M. Rao, *Experts' Perceptions of Autocorrelation: The Hot Hand Fallacy among Professional Basketball Players*, Working Paper, 2009, http://www.justinmrao.com/playersbeliefs.pdf (odczyt: 10.04.2015).

o zależności liniowej. Innymi ważnymi czynnikami wyjaśniającymi skuteczność rzutu są cechy zawodników. Używanie okularów, które zniekształcają widziany obraz, istotnie pogarsza prawdopodobieństwo trafienia do kosza w przypadku młodych koszykarzy. Z uwagi na zespołowy, dynamiczny i kontaktowy charakter gry w koszykówkę wyniki powyższych badań mogą mieć zastosowanie wyłącznie w sytuacji rzutów osobistych z linii rzutów wolnych. Bardzo często w końcowych minutach meczu, gdy piłkę ma przeciwnik, drużyna przegrywająca stosuje taktykę umyślnego piątego faula popełnionego na zawodniku najsłabiej wykonującym rzuty osobiste. Jeżeli prawdopodobieństwo dwukrotnego trafienia z linii rzutów wolnych przez sfaulowanego rywala jest relatywnie niskie, to przejęcie piłki w ten sposób jest korzystne dla drużyny. Zatem rzuty osobiste są ważnym elementem gry, szczególnie w końcowej fazie meczu, gdy jego wynik nie jest jeszcze rozstrzygnięty.

Bibliografia

Albert J., Anderson J.A., *On the Existence of Maximum Likelihood Estimates in Logistic Regression*, „Biometrika” 1984, vol. 71, s. 1–10.

Amemiya T., *Advanced Econometrics*, Harvard University Press, Cambridge 1985.

Breslin G., Hodges N.J., Kennedy R., Hanlon M., Williams A.M., *An especial skill: Support for a learned parameters hypothesis*, „Acta Psychologica” 2010, vol. 134, s. 55–60.

Breslin G., Schmidt R.A., Lee T., *Especial Skills: Generality and Specificity in Motor Learning*, w: *Skill Acquisition in Sport, Research Theory and Practice*, red. A.M. Williams, N. Hodges, Taylor & Francis, London 2012.

Cameron A.C., Trivedi P.L., *Microeconometrics*: *Methods and Application*, Cambridge University Press, New York 2005.

Cramer J.S., *Logit Models From Economics and Other Fields*, Cambridge University Press, Cambridge 2003.

Gourieroux C., *Econometrics of Qualitative Dependent Variables*, Cambridge University Press, Cambridge 2000.

Gruszczyński M., *Model zmiennych jakościowych dwumianowych*, w: *Mikroekonometria: modele i metody analizy danych indywidualnych*, red. M. Gruszczyński, Oficyna a Wolters Kluwer business, Warszawa 2010, s. 53–102.

Hensher D.A., Rose J.M., Greene W.H., *Applied choice analysis: A primer*, Cambridge University Press, Cambridge 2005.

Keetch K.M., Lee T.D., Schmidt R.A., *Especial Skills: Specificity Embedded Within Generality*, „Journal of Sport & Exercise Psychology” 2008, vol. 30(6), s. 723–736.

Keetch K.M., Schmidt R.A., Lee T.D., Young D.E., *Especial Skills: Their Emergence with Massive Amounts of Practice*, „Journal of Experimental Psychology: Human Perception and Performance” 2005, vol. 31, s. 970–978.

Maddala G.S., *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge 1983.

Mikroekonometria: modele i metody analizy danych indywidualnych, red. M. Gruszczyński, Oficyna a Wolters Kluwer business, Warszawa 2010.

Winkelmann R., Boes S., *Analysis of Microdata*, Springer-Verlag, Heidelberg 2006.

Źródła sieciowe

Czyż S.H., Marzec J., Styrkowiec P., *Free throw shot proficiency in basketball – a probability model*, Working Papers, 2012, https://e-uczelnia.uek.krakow.pl/pluginfile.php/68421/mod_page/content/12/Artykuly/2012_Free_throw_shot_proficiency_in_basketball_-_a_probability_model_CzyzMarzecStyrkowiec_ver2012_09.pdf (odczyt: 10.04.2015).

Rao J.M., *Experts' Perceptions of Autocorrelation: The Hot Hand Fallacy among Professional Basketball Players*, Working Paper, 2009, http://www.justinmrao.com/playersbeliefs.pdf (odczyt: 10.04.2015).

* * *

Grouped probit regression analysis for shooting effectiveness in basketball

Summary

In our paper we focused on probit modeling to describe shot effectiveness in basketball free throw. A few hypothesis were verified in repeated experiment. We aimed to determine whether especial skills emerges as a result of other predictors such as age, years of experience, efficiency in shot performance, etc. We also tested a hypothesis regarding visual dependency in especial skills effect. We noticed that some predictors, such as years of training, efficiency in real games, league achievements have strong influence on probability of shot success. Our data confirmed that there is non-linear and negative relationship between shooting distance and shot effectiveness. Blurred vision condition substantially deteriorated shot effectiveness from all shooting distances. However, the shot efficacy at 4,57 m distance was not significantly higher that from any other distances and therefore the hypothesis about context dependency was not confirmed.

Keywords: grouped probit model, especial skills, basketball shooting effectiveness

JEL: C25, C52, Z2

Zgodnie z oświadczeniem autorów, ich udział w powstawaniu artykułu był następujący: Jerzy Marzec – 50%, Stanisław Czyż – 30%, Piotr Styrkowiec – 20%.