

*Agnieszka Bitner*

## **O UŻYTECZNOŚCI METOD STATYSTYCZNYCH W WYCENIE NIERUCHOMOŚCI**

---

### ***ON THE USEFULNESS OF STATISTICAL METHODS IN REAL ESTATE VALUATION***

#### **Streszczenie**

Wykorzystanie metod statystycznych w analizach rynku nieruchomości umożliwia uzyskanie cennych informacji o zachowaniach tego rynku. Informacje te są obiektywne i mogą zostać wykorzystane zarówno do wycen indywidualnych, jak i masowych. W pracy omówiono zastosowanie metod regresyjnych w wycenie nieruchomości, pokazano metody wyznaczania trendu cen nieruchomości, przedstawiono metodę ustalania zależności ceny od pola powierzchni nieruchomości gruntowych, określono postać funkcyjną rozkładu cen jednostkowych nieruchomości gruntowych. Analizę przeprowadzono dla kilkutyśięcnej bazy danych obejmującej transakcje dotyczące nieruchomości gruntowych niezabudowanych dokonane w Krakowie. W artykule dyskutowane są również wybrane problemy związane z zastosowaniem metod regresji wielorakiej w szacowaniu nieruchomości. Są nimi postać funkcyjna modelu, liczba atrybutów uwzględnionych w modelu, trend rynku, skala wartości atrybutów oraz wagi cech. Przedstawione zostały argumenty przemawiające za wyłączeniem daty transakcji z listy cech nieruchomości oraz powody, dla których model powinien zawierać wyraz wolny, a skale atrybutów powinny zaczynać się od zera.

**Słowa kluczowe:** wycenia nieruchomości, model, analiza regresji, atrybuty, wycena masowa

#### ***Summary***

*Statistical methods enable obtaining valuable information about real estate market. The information is objective and can be useful both for individual and mass appraisal. This paper discusses selected issues related to the use of the regression methods in the real estate valuation. They are the method of trend*

*estimation, the method of calculation of the price – size relationship for undeveloped land, and the functional form of the distribution function of unit land prices of undeveloped land. The analysis was based on a few thousands of sale transactions in Kraków. This paper also discusses selected issues related to the use of the multiple linear regression in the real estate valuation. They are the functional form of the model, number of the attributes, trend, scale of the values of the attributes and the form of adjustments. Arguments that the dates of sales should be excluded from the list of real estate attributes are brought forward. Also, the paper argues that the regression model should contain an intercept parameter and the scale of the values of the attributes should start from zero.*

**Key words:** *real estate valuation, model, regression analysis, attributes, mass appraisal*

## WSTĘP

Na rozwiniętych rynkach nieruchomości metody statystyczne stanowią dobrze rozpoznane i często stosowane narzędzie wyceny [Bruce, Sundell 1977; Eckert 1990]. Pierwsza praca, w której wartość nieruchomości była estymowana za pomocą regresji wielorakiej powstała już na początku ubiegłego wieku [Haas 1922]. Zastosowanie komputerów umożliwiło nie tylko przechowywanie dużych baz danych, łatwy i szybki dostęp do danych ale również usprawniło sam proces wyceny. Duże możliwości stwarza również wykorzystanie prężnie rozwijającego się systemu GIS (*Geographic Information System*), umożliwiającego na przykład szybką przestrzenną identyfikację nieruchomości. Połączenie informacji o nieruchomościach z systemem GIS znacznie ułatwiło pracę rzeczoznawców majątkowych na tych rynkach.

Obecnie w środowisku polskich rzeczoznawców majątkowych nie ma już wątpliwości dotyczących konieczności zastosowania metod statystycznych w wycenie [Hozer i in. 2002; Żróbek, Bełej 2000; Grzesik i in. 2007; Sawiłow 2010]. Dostępność danych i łatwość ich pozyskania powodują, że każdy z rzeczoznawców majątkowych dysponuje setkami, jeśli nie tysiącami danych transakcyjnych z danego rynku lokalnego. Wykorzystanie metod statystycznych w analizach rynku nieruchomości umożliwia uzyskanie cennych informacji o zachowaniach tego rynku zawartych w dużej liczbie obserwacji. Informacje te są niedostępne, jeśli każdą z nieruchomości będziemy traktowali indywidualnie. Otrzymane wyniki analiz statystycznych są obiektywne i mogą zostać wykorzystane na innych podobnych rynkach lokalnych. Ze względu na konieczność posiadania dużej liczby danych transakcyjnych, metody statystyczne najbardziej nadają się do wyceny nieruchomości typowych [Cellmer 1999] oraz przy masowej wycenie nieruchomości [Mark, Goldberg 1988]. Wówczas powodują one zobiektywizowanie, usprawnienie oraz obniżenie kosztów wyceny. W przypadku wycen indywidualnych wyniki otrzymane z zastosowaniem metod statystycznych stanowią obiektywny punkt odniesienia wolny od uznaniowości wyceniającego.

Celem pracy jest pokazanie użyteczności metod statystycznych w wycenie nieruchomości. Na podstawie analizy statystycznej rynku nieruchomości gruntowych niezabudowanych określę metodę wyznaczania trendu cen nieruchomości, przedstawię metodę ustalania zależności ceny od pola powierzchni nieruchomości gruntowych, określę postać funkcyjną rozkładu cen jednostkowych nieruchomości gruntowych oraz omówię zastosowanie metody regresji wielorakiej w szacowaniu wartości nieruchomości.

### **BAZA DANYCH**

Analizę przeprowadzono dla kilkutysięcznej bazy danych. Dane pochodzą z aktów notarialnych dotyczących transakcji kupna-sprzedaży działek niezabudowanych położonych w granicach administracyjnych Krakowa. Badany okres czasu obejmuje cztery lata, od 1.01.1996 r. do 31.12.1999 r. Baza składa się z 6505 rekordów. Z bazy usunięto transakcje dokonane w ramach przetargu oraz działki, dla których wydane zostało pozwolenie na budowę. Poza tym nie dokonano żadnych innych eliminacji. W szczególności nie usunięto transakcji, w których cena za 1 m<sup>2</sup> gruntu była – w stosunku do średniej – bardzo niska lub wysoka. Nie istnieją bowiem obiektywne kryteria, zgodnie z którymi można by dokonać takiej eliminacji. Każda transakcja wnosi swoją informację i wpływa na zachowanie rynku. Trudno jest zatem zdecydować z góry o tym, która cena jest „lepsza”, a która „gorsza”. Celem moich badań było bowiem wypracowanie uniwersalnych metod pomocnych przy wartościowaniu gruntów. W celu wyeliminowania wpływu przeznaczenia gruntu na badane zależności statystyczne, analizę wykonano dla gruntów przeznaczonych pod niską zabudowę mieszkaniową, oznaczonych jako M4. Są one najczęstszym przedmiotem obrotu i stanowią 37,23% transakcji, są więc najbardziej reprezentatywne. W analizie nie uwzględniono transakcji dotyczących nieruchomości o szczególnych sposobach użytkowania: sprzedanych jako dopełnienie sąsiedniej nieruchomości, działki drogi, działki wydzielone pod rurociąg gazowy, działki z linią wysokiego napięcia. Ostatecznie analizą objęto 1777 transakcji.

### **TREND PROCENTOWYCH ZMIAN CEN**

Analizę statystyczną rynku nieruchomości rozpoczęto od wyznaczenia trendu procentowych zmian cen jednostkowych nieruchomości gruntowych o przeznaczeniu M4 w badanym okresie. Wskaźnik trendu zmian cen nieruchomości gruntowych został obliczony metodą regresji liniowej. Jest on współczynnikiem kierunkowym prostej regresji, dopasowanej do danych, metodą najmniejszych kwadratów (rys. 1). Otrzymany w ten sposób wskaźnik służy do korekty cen sprzedaży nieruchomości ze względu na czas, jaki upłynął od daty

transakcji do momentu szacowania nieruchomości. W ogólnym przypadku, jeśli dysponujemy wskaźnikiem miesięcznego procentowego wzrostu cen  $T$ , procentowej korekty cen dokonuje się zgodnie ze wzorem:

$$C_{sp} = C_t \times (1 + n \times T), \quad (1)$$

gdzie:

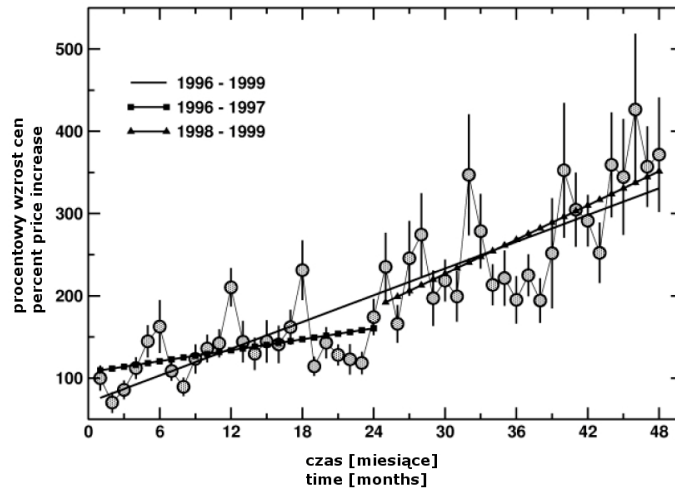
- $C_{sp}$  – cena sprowadzona,
- $C_t$  – cena transakcyjna,
- $n$  – liczba miesięcy jaka upłynęła od daty transakcji do określonej daty wyceny. Jeśli wskaźnik  $T$  był znacząco różny w kolejnych przedziałach czasowych badanego okresu, korekta cen powinna uwzględnić ten fakt i odbywać się kilkietapowo, zgodnie ze wzorem:

$$C_{sp} = C_t \times (1 + n_1 \times T_1) \times (1 + n_2 \times T_2) \times \dots \times (1 + n_k \times T_k), \quad (2)$$

gdzie:

- $n_i$  – liczba miesięcy w  $i$ -tym przedziale czasowym ( $i = 1, \dots, k$ ),
- $T_i$  – wskaźnik miesięcznego procentowego wzrostu cen w  $i$ -tym przedziale czasowym,
- $k$  – liczba przedziałów, na jakie podzielono badany okres czasu.

Punkty na rysunku 1 przedstawiają ceny transakcyjne uśrednione w kolejnych miesiącach. Skala na osi rzędnych podaje procentową zmianę przeciętnej ceny jednostkowej wyznaczonej dla poszczególnych miesięcy w stosunku do średniej ceny w pierwszym miesiącu badanego okresu. Dane zostały przedstawione na wykresie wraz z błędami. W badanym czteroletnim okresie czasu widać wyraźną zmianę w zachowaniach rynku nieruchomości gruntowych na przełomie lat 1997 i 1998. Z tego powodu estymowane linie trendów wyznaczono dla trzech okresów: 1996–1997, 1998–1999 oraz 1996–1999. Na podstawie danych 1996–1999 współczynnik regresji liniowej jest szacowany na poziomie  $T = 5,42$  [Bitner 2001, 2003]. Oznacza to, że z każdym miesiącem badanego okresu średni wzrost cen nieruchomości gruntowych o przeznaczeniu M4 wynosił 5,42%. Średni błąd statystyczny tego oszacowania wynosi 0,49%. Wskaźniki trendu dla kolejnych dwuletnich okresów wynoszą odpowiednio 2,22 oraz 6,93% na miesiąc. Wynika stąd, że właściwą metodą dokonania korekty cen dla czteroletniego okresu jest dwuetapowe sprowadzenie cen. Rysunek 1 dostarcza jeszcze innych ważnych informacji. Po pierwsze, ekstrapolowanie cen na podstawie nawet dwuletniego trendu jest dość ryzykowne. Po drugie, widoczna jest cykliczność w zachowaniu się cen nieruchomości gruntowych o przeznaczeniu M4 w Krakowie (szczególnie wyraźna dla okresu 1996–1997). Największe wzrosty cen obserwuje się średnio co pół roku.



Źródło: Bitner 2003.

**Rysunek 1.** Linie trendu procentowych zmian cen wyznaczone dla miesięcznych okresów uśredniania. Wskaźnik trendu rynku,  $T$ , jest współczynnikiem kierunkowym prostej regresji

**Figure 1.** Trend lines. The points represents average monthly data. The trend coefficient,  $T$ , is the slope of the regression line

### ZALEŻNOŚĆ CENA – POLE POWIERZCHNI

Pole powierzchni nieruchomości gruntowej jest jedną z podstawowych cech uwzględnianych w każdym podejściu do szacowania gruntów. Obiektywne określenie współczynnika przeliczeniowego służącego do oceny wpływu wielkości pola powierzchni na cenę uprościłoby proces wyceny gruntów. Wyznaczenie poprawki w postaci procentowej (a nie kwotowej) umożliwia zastosowanie otrzymanych wyników dla podobnych rynków lokalnych. Zależność ceny jednostkowej nieruchomości gruntowej od wielkości pola powierzchni intuicyjnie wydaje się oczywista: za większe powierzchnie powinniśmy zapłacić mniej w przeliczeniu na jednostkę. Zagadnienie to było dyskutowane pod względem jakościowym [Żróbek, Belej 2000] oraz ilościowym [Prystupa 2001].

Zasadnicza różnica między prezentowanym tutaj i szczegółowo opisanym w pracy [Bitner 2008] sposobem wyznaczania poprawki na pole powierzchni, a metodami spotykanymi w literaturze polega na sposobie doboru danych. Mianowicie, aby określić wspomnianą poprawkę zazwyczaj dokonuje się selekcji nieruchomości o bardzo zbliżonych wartościach pozostałych cech i na ich pod-

stawie wyznacza zależność cena – pole. Wymaga to wiedzy i doświadczenia wyceniającego. Zakłada się przy tym, że nieruchomości różnią się między sobą tylko wartością atrybutu „pole powierzchni”. Założenie to czyniono między innymi w metodzie porównywania parami [Prystupa 2001; Źróbek, Belej 2000] oraz w analizie regresji [Hozer 1999]. W praktyce znalezienie nawet kilku nieruchomości różniących się jedną cechą (polem powierzchni) jest bardzo trudne. W prezentowanej metodzie, dzięki uwzględnieniu wszystkich transakcji występujących na rynku, wykorzystuje się całą dostępną ilość informacji o zależności ceny – pola. Dane do analizy nie były dobierane, obliczenia wykonano wykorzystując wszystkie transakcje rynkowe dotyczące gruntów o przeznaczeniu M4. Jest to metoda bardziej czasochłonna od metod prezentowanych w cytowanych pracach, ale jest ona wolna od subiektywnego doboru danych bazowych. Ponadto metodę tę można zalgorytmizować, co ma znaczenie zwłaszcza przy masowej wycenie.

**Model zależności ceny jednostkowej od wielkości pola powierzchni.**

Rysunek 2a przedstawia ogólny schemat wyznaczania poprawki procentowej w modelu krzywoliniowym. Wyznaczana jest ona w następujący sposób: Niech  $C(a_1, \dots, a_n, S)$  oznacza cenę jednostkową wybranej nieruchomości w funkcji atrybutów  $\{a_1, \dots, a_n, S\}$ . Ponieważ zakładamy, że całą zależność ceny jednostkowej od pola powierzchni  $S$  opisuje funkcja  $\delta(S)$ , wpływ pozostałych atrybutów uwzględniamy w cenie  $C_0$ , stąd cenę  $C$  nieruchomości możemy zapisać w postaci iloczynu:

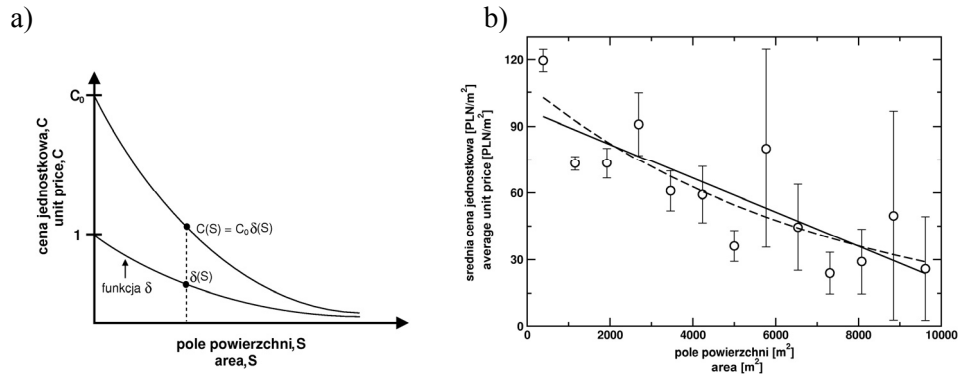
$$C(S) = C_0(a_1, \dots, a_n) \cdot \delta(S). \quad (3)$$

Poprawka  $\delta$  jest zatem funkcją malejącą o dodatnich wartościach. Dodatkowo, w celu nadania funkcji  $\delta$  naturalnej interpretacji jako względnego spadku ceny, przyjmuje się, że  $\delta(0) = 1$ .

Nieznana cenę jednostkową  $C(S')$  nieruchomości o polu powierzchni  $S'$  wyznaczoną na podstawie znanej ceny,  $C(S)$ , nieruchomości o polu  $S$  i tych samych (bardzo zbliżonych) wartościach pozostałych atrybutów możemy wyrowadzić ze wzoru:

$$C(S') = (\delta(S')/\delta(S)) \cdot C(S). \quad (4)$$

Podczas analizowania rynku musimy założyć konkretną postać funkcji  $\delta$ . Wyznaczając zależność cena – pole, rozważono dwa modele: krzywoliniowy i liniowy (Rys. 2b). Najprostszą funkcją krzywoliniową, która pod względem matematycznym odpowiada badanej zależności jest funkcja wykładnicza,  $\delta(S) = \exp(-\alpha_e \cdot S)$ . Dla modelu liniowego  $\delta(S) = 1 - \alpha_l \cdot S$ .



Źródło: Bitner 2008.

**Rysunek 2.** (a) Zależność ceny jednostkowej,  $C$ , w funkcji pola powierzchni  $S$ . Funkcja  $\delta$  odpowiada hipotetycznej sytuacji, gdy cena  $C_0$  wynosi 1 PLN za  $m^2$ ; (b) Zależność średniej ceny jednostkowej nieruchomości gruntowej od wielkości pola powierzchni, przy podziale na 13 przedziałów. Punkty na wykresie podane zostały wraz z błędami statystycznymi i przedstawiają ceny jednostkowe uśrednione w poszczególnych przedziałach. Do punktów dopasowano krzywe w modelach wykładniczym (linia przerywana) i liniowym (linia ciągła)

**Figure 2.** (a) The relationship between the unit price,  $C$ , and the parcel's size  $S$ .  $\delta$  function corresponds to the hypothetical situation when the unit price  $C_0$  is 1 PLN; (b) The relationship between the average unit price and a lot size, the number of bins is 13. The points in the graph were plotted with statistical errors and represent average unit prices in each bin. The two curves are exponential (dashed line) and linear (solid line) fits to the data

**Metoda uśredniania.** Odmienność prezentowanej metody wyznaczenia funkcji  $\delta$ , wykorzystującej informacje zawarte w dużych bazach danych, polega po pierwsze na wykorzystaniu wszystkich wolnorynkowych danych transakcyjnych (brak uznaniowości przy doborze danych). Po drugie na uśrednieniu cen w poszczególnych przedziałach klasowych. Uśrednienie cen wewnątrz przedziału klasowego stanowi zasadniczy element prezentowanej metody wyznaczenia funkcji  $\delta$ . Chodzi o to, by w poszczególnych przedziałach znalazły się nieruchomości, dla których pozostałe czynniki cenotwórcze przyjmują jak najwięcej różnych wartości. Różnice w wartościach atrybutów znoszą się w dużym stopniu przy opisanym wyżej uśrednianiu. Dzięki temu wartości pozostałych atrybutów zostały ujednolicone. Różnorodność wartości atrybutów będąca wadą w przypadku małych baz danych w prezentowanej metodzie jest wielką zaletą.

**Wyniki analizy.** Analizę wykonano dla nieruchomości o przeznaczeniu M4 o powierzchniach do  $100 m^2$ . Ograniczenie zakresu powierzchni było spowodowane małą liczbą nieruchomości transakcyjnych o większych powierzchniach. Ceny sprowadzono na dzień ostatniej transakcji w bazie stosując wskaź-

nik  $T$ . Obliczenia wykonano dla podziału na 13 (rys. 3b) oraz 5 klas równej długości. Do otrzymanych punktów dopasowano metodą najmniejszych kwadratów funkcje wykładniczą i liniową. Dla *wykładniczej* postaci modelu, przy podziale na 13 oraz 5 klas, otrzymane współczynniki kierunkowe wynoszą odpowiednio  $\alpha_e^{(13)} = 1,37 \cdot 10^{-4} \text{ m}^{-2}$  oraz  $\alpha_e^{(5)} = 1,40 \cdot 10^{-4} \text{ m}^{-2}$ . Różnice parametrów funkcji  $\delta$  przy zastosowaniu różnych podziałów na klasy są zatem niewielkie. W praktyce można przyjąć wartość uśrednioną  $\alpha_e = 1,38 \cdot 10^{-4} \text{ m}^{-2}$ . Dla *liniowej* postaci modelu, przy podziale na 13 oraz 5 klas, otrzymane współczynniki wynoszą odpowiednio  $\alpha_l^{(13)} = 7,87 \cdot 10^{-5} \text{ m}^{-2}$  oraz  $\alpha_l^{(5)} = 8,06 \cdot 10^{-5} \text{ m}^{-2}$ . Otrzymane wartości współczynników kierunkowych prostych przy podziale na 13 i 5 klas są zatem bardzo zbliżone. W zastosowaniach można zatem przyjąć wartość uśrednioną  $\alpha_l = 7,97 \cdot 10^{-5} \text{ m}^{-2}$ . Interpretując otrzymane dla modelu liniowego wyniki stwierdzamy, że wzrost powierzchni nieruchomości o jeden ar powoduje spadek ceny jednostkowej o około 0,8%.

### ROZKŁAD JEDNOSTKOWYCH CEN

Określenie postaci funkcyjnej rozkładu jednostkowych cen transakcyjnych nieruchomości gruntowych dostarcza istotnych informacji o rynku nieruchomości jak też o jego uczestnikach [Bitner 2009]. Znajomość postaci funkcyjnej rozkładu cen jest potrzebna rzeczoznawcom, ponieważ stosowane w wycenie testy statystyczne zakładają często rozkład normalny cechy. Z kształtem rozkładu wiąże się również inna ważna kwestia mająca duże znaczenie przy wycenie nieruchomości. Mianowicie, czy wartość średnia jest dobrym reprezentantem cen w analizie rynku. Dla rozkładów niesymetrycznych dokładniejszą informację o poziomie cen występujących na rynku daje wartość modalna. Histogramy przedstawione w pracy [Hozer i in. 2002] wskazują, że rozkład cen nie zawsze jest symetryczny.

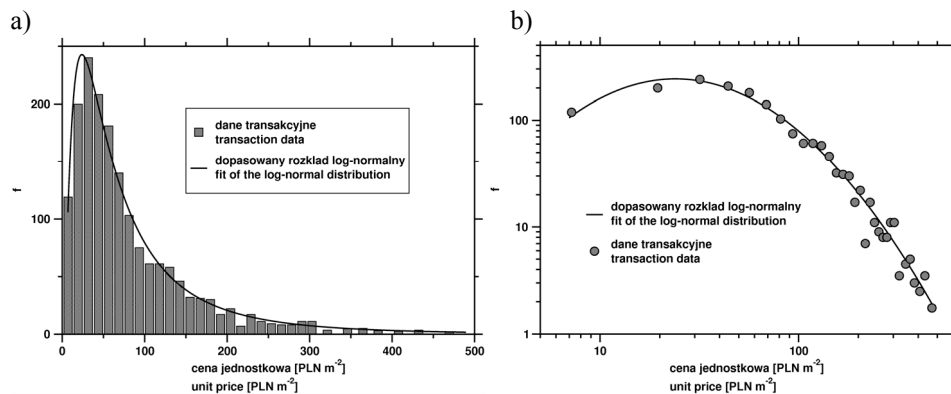
**Wyniki analizy.** Wyjściowy zbiór 1777 danych transakcyjnych obejmował ceny jednostkowe,  $c$ , zmieniające się w przedziale od  $c_{\min} = 1,03 \text{ PLN m}^{-2}$  do  $c_{\max} = 1428,83 \text{ PLN m}^{-2}$ . Do analizy wykorzystano 1752 transakcje, w których ceny jednostkowe nie przekraczały  $500 \text{ PLN m}^{-2}$ . Pozostałe 25 transakcji – stanowiące około 1% bazy – usunięto. Celem analizy było przetestowanie hipotezy  $H_0$ , że rozkład  $f(c)$  jest opisany przez rozkład log-normalny (LN), dany wzorem:

$$f(c) = \frac{1}{c\sigma\sqrt{2\pi}} \exp\left(-\frac{\ln^2(c/\mu)}{2\sigma^2}\right). \quad (5)$$

Rozkład LN jest określony tylko dla dodatnich wartości zmiennej  $c$ , i jest rozkładem asymetrycznym, skośnym prawostronnie. Rozkład ten charakteryzują średnia,  $\mu$ , oraz odchylenie standardowe,  $\sigma$ , zmiennej  $\ln(c)$ . W celu zweryfikowania hipotezy  $H_0$  przeprowadzono test chi-kwadrat Pearsona.



Rysunek 3a przedstawia histogram  $f(c)$  cen jednostkowych nieruchomości gruntowych, otrzymany dla podziału na  $n = 32$  przedziały. Linia ciągłą wykreślona została dystrybucja LN z parametrami wyznaczonymi przez dopasowanie metodą najmniejszych kwadratów. Jakość dopasowania funkcji rozkładu do danych ocenić można lepiej na wykresie podwójnie logarytmicznym. Jak widać na Rys. 3b, dane empiryczne bardzo dobrze pasują do dystrybucji LN, która na skali podwójnie logarytmicznej ma kształt paraboli.



Źródło: Bitner 2009.

**Rysunek 3.** (a) Histogram jednostkowych cen nieruchomości gruntowych otrzymany dla  $n = 32$  przedziałów. Linia ciągła przedstawia rozkład LN z parametrami wyznaczonymi przez dopasowanie metodą najmniejszych kwadratów;  
(b) Te same dane na wykresie podwójnie logarytmicznym

**Figure 3.** (a) Histogram of the unit land prices obtained for  $n = 32$  bins. The solid line represents the LN distribution fitted to the data using the least-squares method;  
(b) Same as in (a) plotted in the double logarithmic scale

Dla podziału na  $n = 32$  przedziały wartości parametrów wynoszą  $\mu = 59,8 \pm 7,9$  PLN  $m^{-2}$  i  $\sigma = 0,956 \pm 0,080$ . Wartość statystyki  $\chi^2 = 30,857$ , wartość krytyczna dla poziomu zgodności  $\alpha = 0,1$  wynosi  $\chi^2_{n-4,\alpha} = 42,585$ . Wartość statystyki  $\chi^2$  jest zatem mniejsza od wartości krytycznej. Oznacza to, że nie ma podstaw do odrzucenia postawionej hipotezy  $H_0$ .

Przeprowadzona analiza bazuje na danych z lat 1996–1999. Należy jednak podkreślić, że, mimo, iż są to dane sprzed prawie dziesięciu lat, nie ma to znaczenia dla prezentowanego wniosku. Dotyczy on bowiem postaci funkcyjnej rozkładu cen nieruchomości. Należy oczekiwać, że postać tego rozkładu jest niezależną od czasu własnością rynku. Z biegiem lat zmieniać mogą się jedynie parametry rozkładu – wielkości  $\mu$  oraz  $\sigma$ .

## MODEL REGRESJI WIELORAKIEJ W SZACOWANIU NIERUCHOMOŚCI

Celem niniejszego rozdziału jest zasygnalizowanie kluczowych problemów związanych ze stosowaniem metod liniowej regresji wielorakiej w szacowaniu nieruchomości. Problemy te dotyczą: postaci funkcyjnej modelu, trendu rynku, liczby atrybutów, skali wartości atrybutów, wag cech. Sposoby oceny statystycznej zaproponowanego modelu wraz z przykładem weryfikacji statystycznej wyceny nieruchomości metodą analizy statystycznej rynku opisano w pracy [Bitner 2007].

**Postać funkcyjna modelu.** Dobrze skonstruowany model statystyczny jest kompromisem między dwoma skrajnościami: nadmiernym uproszczeniem, a nadmiernym nagromadzeniem szczegółów. Dokonując wyboru postaci funkcyjnej modelu, powinno się zatem koncentrować na najprostszych formułach matematycznych [Rao, Toutenburg 2003; Aczel 2000]. W przypadku szacowania wartości nieruchomości na podstawie zbioru nieruchomości reprezentatywnych najbardziej odpowiednimi wydają się modele liniowe. Zmienną objaśnianą w modelu jest wartość wycenianej nieruchomości, atrybuty są zmiennymi objaśniającymi. Stosowanie modeli liniowych w wycenie jest uzasadnione z dwóch powodów: (1) jest to najprostszy sposób określenia wpływu atrybutów na cenę, i najbardziej właściwy w sytuacji kiedy nie jest znana funkcyjna postać tej zależności. (2) niemal wszystkie funkcje dają się w pewnym zakresie linearyzować.

**Wyraz wolny.** Dokonując konstrukcji modelu, zakładamy, że jednostkowe ceny nieruchomości w bazie,  $C^{(1)}, \dots, C^{(n)}$ , są liniową funkcją atrybutów  $X_1, \dots, X_m$ , na badanym rynku lokalnym. Model wartości nieruchomości powinien opisywać poprawki względem średniej ceny nieruchomości na rynku,  $\bar{C}$ . Nawet teoretycznie model powinien wykluczyć sytuację, kiedy wartość wycenianej nieruchomości nie jest dodatnia. Powyższe warunki możemy spełnić odpowiednio definiując skalę atrybutów oraz wprowadzeniu do modelu wyrazu wolnego  $a_0$ . Skala każdego atrybutu powinna zaczynać się od zera. Postać funkcyjna tego modelu jest wtedy następująca:

$$W = a_0 + a_1 X_1 + \dots + a_m X_m, \quad (6)$$

gdzie:

- $a_i$  – współczynniki wyznaczone metodą najmniejszych kwadratów,
- $W$  – wartość 1 m<sup>2</sup> nieruchomości wycenianej. Wyraz wolny ma tutaj prostą interpretację jako wartość 1 m<sup>2</sup> hipotetycznej „najgorszej” nieruchomości. Wówczas wartość nieruchomości  $W = a_0$ . Każda, nawet najgorsza nieruchomość ma wartość dlatego  $a_0$  powinna być dodatnia. W przeciwnym wypadku model jest błędny i nie spełnia przyjętych założeń. W tej sytuacji należy jeszcze raz zweryfikować bazę nieruchomości reprezentatywnych, sprawdzić czy w modelu uwzględniliśmy wszystkie istotne atrybuty.

**Liczba zmiennych objaśniających (atrybutów) w modelu.** Ważną częścią konstrukcji modelu wyceny jest ustalenie liczby atrybutów jakie należy uwzględnić. Jedną ze wskazówek, którą możemy kierować się przy ustalaniu liczby atrybutów daje nam psychologia. Mianowicie, wyniki badań psychologicznych wykazały, że człowiek nie potrafi podejmować decyzji na podstawie więcej niż około siedmiu kryteriów [Tomaszewski 1975]. Podobnie, nie potrafi właściwie wartościować cech według ich ważności w liczbie przekraczającej siedem pozycji. Zatem, zgodnie z tą wskazówką, liczba atrybutów nie powinna raczej przekraczać kilku. Oznacza to, że niezbędna jest selekcja (redukcja) atrybutów [Eckert 1990]. Dokonując redukcji zbioru wstępnie wytypowanych zmiennych objaśniających, należy przestrzegać zasady, iż w modelu powinny się znaleźć zmienne silnie skorelowane ze zmienną objaśnianą (atrybuty mające istotny wpływ na wartość nieruchomości) i jednocześnie bardzo słabo skorelowane między sobą. Skorelowane zmienne objaśniające pozbawiają się nawzajem mocy wyjaśniającej.

Jeśli przyjmiemy zbyt wiele atrybutów  $X_i$  (np. kilkanaście), wtedy na pewno niektóre z nich będą, w zbiorze wyników obserwacji, z sobą skorelowane. Eliminując jedną ze zmiennych skorelowanych z modelu, znacznie go poprawimy, tracąc przy tym niewiele informacji o zmiennej objaśnianej. Jeśli zmienne są skorelowane to część informacji, którą o zmiennej objaśnianej niesie jedna z tych zmiennych, jest też zawarta w drugiej. Idealną sytuacją byłoby, gdyby atrybuty nie były ze sobą skorelowane (brak współliniowości). Każda cecha cenotwórcza przekazywałaby wtedy swoją część informacji o wartości wycenianej nieruchomości – informacji, której nie przekazywałaby żadna inna cecha. Podstawową miarą współliniowości dwóch atrybutów jest współczynnik korelacji.

**Trend rynku.** Trend to wielkość, której wpływ na cenę jesteśmy w stanie wyznaczyć niezależnie od pozostałych atrybutów (wskaźnik  $T$ ). Należy to uczynić z następujących powodów: (1) „Stan rynku” (trend) jest szczególnym atrybutem, zaliczanym w literaturze [Eckert 1990] do *nonproperty characteristics* w odróżnieniu od takich atrybutów jak, na przykład, lokalizacja czy pole powierzchni, zaliczanych do *property characteristics*. Trend jest zatem właściwością rynku a nie cechą nieruchomości. Z tego powodu w analizie powinny być używane sprowadzone ceny transakcyjne. (2) Trendu wzrostu cen wyznaczony niezależnie jest znacznie dokładniejszy niż w przypadku, gdy traktuje się go jako jedną ze zmiennych w modelu regresji. Do wyznaczenia trendu rynku możemy bowiem wykorzystać wszystkie transakcje rynkowe dotyczące nieruchomości danego typu na rynku lokalnym. (3) Dzięki wyznaczeniu trendu niezależnie model wyceny zostanie odciążony, ponieważ zmniejszona zostanie liczba zmiennych.

Do poprawnego wyznaczenia trendu trzeba dysponować odpowiednio dużą bazą nieruchomości reprezentatywnych zbieranych w dłuższym okresie czasu (minimum pół roku). Najbardziej właściwym do sprowadzenia cen na jedną datę

jest wskaźnik  $T$  wyznaczony metodą regresji liniowej na podstawie bazy cen transakcyjnych [Bitner 2001, 2003].

**Skala wartości atrybutów oraz wyznaczenie wag cech.** Atrybuty nieruchomości mogą przyjmować wartości ciągle lub dyskretne. Na skalach dyskretnych wartościuje się cechy o charakterze jakościowym lub dychotomicznym. Obie skale można zawsze przekształcić na skalę ciągłą z przedziału  $[0, 1]$ . W przypadku skali ciągłej można to osiągnąć poprzez wydzielenie wartości cechy przez rozsądnie ustaloną wartość maksymalną. Dla skali dyskretnej, dzieląc wartość atrybutu przez maksymalną wartość cechy. Wprowadzenie ciągłej skali znormalizowanej  $[0, 1]$  jest wskazane z następujących powodów: (1) współczynniki,  $a_1, \dots, a_m$ , opisują wtedy *kwotową* zmianę wartości nieruchomości odpowiadającą przyrostowi wartości zmiennej  $X_i$ , (2) możliwe jest określenie wartości „najgorszej“ nieruchomości, dla której każda ze zmiennych przyjmuje wartość zero. (3) przeskalowanie wartości zmiennych  $X_i$  nie ma wpływu na wartość współczynnika  $a_0$ , a jedynie na wartości współczynników kierunkowych.

Współczynniki  $a_1, \dots, a_m$  nie są wagami cech w rozumieniu standardu III.7, ponieważ ich suma nie musi równać się jedności oraz ich wartości mogą być ujemne. Wagi cech możemy w prosty sposób wyznaczyć ze współczynników  $a_i$  za pomocą następujących wzorów:  $w_i = |a_i|/\sum|a_i|$ ,  $i = 1, \dots, m$ . Wagi  $w_i$  opisują względny wpływ atrybutów na cenę. Współczynnik  $a_0$  mówi jak silnie wartość nieruchomości zależy od przyjętych w modelu atrybutów. Jeśli wartość  $a_0$  jest znacznie większa od wartości bezwzględnych pozostałych współczynników oznacza to, że przyjęte atrybuty nieruchomości w niewielkim stopniu wpływają na jej wartość.

## PODSUMOWANIE

Celem niniejszego artykułu było zwrócenie uwagi na użyteczność metod statystycznych w szacowaniu wartości nieruchomości. W przypadku wyceny masowej stanowią one podstawowe narzędzie analizy. Zastosowanie metod statystycznych do analizy rynku w przypadku wycen indywidualnych dostarcza obiektywnych informacji o stanie rynku i zachodzących na nim zależnościach. Informacje te mogą zostać również wykorzystane na podobnych rynkach lokalnych.

Analizę statystyczną rynku rozpoczęto od wyznaczenia wskaźnika zmian cen w badanym okresie czasu. Stosując regresję liniową wyznaczono wskaźnik trendu rynku, który wykorzystano do sprowadzenia cen na datę ostatniej transakcji w bazie danych. Sprowadzone dane transakcyjne posłużyły następnie do zbadania wpływu wielkości pola powierzchni na cenę nieruchomości gruntowych oraz określenia rozkładu jednostkowych cen nieruchomości gruntowych.

Specyfika zaprezentowanej metody określania zależności cena – pole polega *po pierwsze* na wykorzystaniu wszystkich transakcji, czyli całej dostępnej informacji o wpływie pola powierzchni na cenę na danym rynku lokalnym. Na żadnym etapie analizy nie jest potrzebny subiektywny wybór danych bazowych. *Po drugie* na uśrednieniu cen w poszczególnych przedziałach klasowych, na jakie został podzielony cały zakres pól powierzchni w bazie. Dzięki uśrednieniu cen, wartości pozostałych atrybutów zostały ujednoczone, a zatem zostało spełnione podstawowe założenie o podobieństwie wartości pozostałych cech. Nieruchomości reprezentujące poszczególne przedziały różnią się zatem tylko jedną cechą – polem powierzchni. Metoda uśredniania umożliwia obiektywne wyznaczenie zależności cena – pole w postaci poprawki procentowej, która może być wykorzystana na innych rynkach lokalnych.

Wykazano, że rozkład jednostkowych cen transakcyjnych nieruchomości gruntowych jest zgodny z rozkładem log-normalnym. Otrzymany wynik wskazuje na potrzebę dokładnej wstępnej analizy statystycznej danych o nieruchomościach. Zakładając *a priori*, że rozkład cen transakcyjnych jest rozkładem Gaussa, można popełnić błąd. W szczególności, takie nie zweryfikowane założenie powoduje stosowanie nieodpowiednich testów statystycznych. Dlatego by poprawnie dobrać i stosować narzędzia statystyczne warto zbadać postać funkcyjną rozkładu cen.

W pracy przeprowadzono również dyskusję wybranych problemów związanych z wykorzystaniem modelu regresji wielorakiej w szacowaniu wartości nieruchomości. Przedstawiono argumenty przemawiające za wyłączeniem daty transakcji z listy cech nieruchomości. Trend jest bowiem właściwością rynku a nie cechą nieruchomości. Z tego powodu w analizie powinny być używane sprowadzone ceny transakcyjne. Trend wzrostu cen wyznaczony niezależnie jest znacznie dokładniejszy niż w przypadku, gdy traktuje się go jako jedną ze zmiennych w modelu regresji wielorakiej. Dodatkowo, dzięki wyłączeniu dat transakcji model wyceny zostanie odciążony, ponieważ zmniejszona zostanie liczba jego parametrów. Ponadto przedyskutowany został problem uwzględnienia wyrazu wolnego w modelu oraz wyboru skal wartości atrybutów. Przedstawiono powody, dla których model powinien zawierać wyraz wolny, a skale atrybutów powinny zaczynać się od zera. Najważniejszym z nich jest fakt, że każda nieruchomość ma wartość a model wyceny – nawet teoretycznie – powinien wykluczyć sytuację, kiedy wartość nieruchomości nie jest dodatnia. Wyraz wolny ma prostą interpretację jako wartość 1 m<sup>2</sup> hipotetycznej „najgorszej” nieruchomości.

## BIBLIOGRAFIA

- Aczel A. D. *Statystyka w zarządzaniu*. PWN, Warszawa 2000.
- Bitner A. *Sprowadzanie cen na określoną datę*. Świat Nieruchomości 53, 2001, s. 721–726.
- Bitner A. *Wskaźniki wzrostu cen nieruchomości gruntowych dla dużych rynków lokalnych*. Materiały III Konferencji Naukowo-Technicznej PSRWN, Gdańsk 2003.
- Bitner A. *Konstrukcja modelu regresji wielorakiej przy wycenie nieruchomości*. Acta Sci. Pol., Administratio Locorum, 6(4), 2007, s. 59–66.
- Bitner A. *Zależność cena – pole powierzchni dla nieruchomości gruntowych niezabudowanych*. Acta Sci. Pol., Administratio Locorum, 7(1), 2008, s. 41–53.
- Bitner A. *Rozkład jednostkowych cen nieruchomości gruntowych*. Acta Sci. Pol., Administratio Locorum, 8(4), Olsztyn 2009, s. 41–50.
- Bruce R. W., Sundell D. J. *Multiple regression analysis: history and applications in the appraisal profession*. Real Estate Appraiser, Jan/Feb, 1977, s. 37–44.
- Cellmer R. *Propozycja procedury wyceny nieruchomości metodą analizy statystycznej rynku*. Wycena Nr 6 (47). 1999.
- Grzesik K., Żróbek R., Żróbek S. *Kierunki standaryzacji zasad wyceny nieruchomości w warunkach globalizacji [w:] Wpływ globalizacji na rozwój rynków nieruchomości*. Wyd. TNN, 15, 2007, s. 9–22.
- Eckert J. K. (editor). *Property Appraisal and Assessment Administration*. International Association of Assessing Officers, Chicago 1990.
- Haas G. C. *Sales Prices as a Basic for Farm Land Appraisal*. Technical Bulletin 9, Agricultural Experimental Station, The University of Minnesota, St. Paul 1922.
- Hozer J. *Wykorzystanie wyników analiz statystycznych rynku nieruchomości do szacowania nieruchomości*. Zesz. Nauk. Uniw. Szczecińskiego Nr 233, Prace Katedry Ekonometrii i Statystyki, 1999.
- Hozer J., Kokot S., Kuźmiński W. *Metody analizy statystycznej rynku w wycenie nieruchomości*. PFSRM, Warszawa 2002.
- Mark J., Goldberg M. A. *Multiple regression analysis and mass assessment*. A review of the issues, Appraisal Journal 56, 1988, s. 89–109.
- Prystupa M. *Wycena nieruchomości przy zastosowaniu podejścia porównawczego*. PFSRM, Warszawa 2001.
- Rao C. R. Toutenburg H. *Linear Models*. Springer-Verlag, New York 1995.
- Sawilow E. *Problematyka określania nieruchomości metodą analizy statystycznej rynku*. Wyd. TNN, 18, 1, 2010, s. 21–30.
- Standard III.7*. PFSRM, Warszawa 2003.
- Tomaszewski T. (red.). *Psychologia*. Wyd. 5, PWN Warszawa 1975.
- Żróbek S., Belej M. *Podejście porównawcze w szacowaniu nieruchomości*. Educaterra, Olsztyn 2000.

Dr Agnieszka Bitner  
Katedra Zastosowań Matematyki  
Uniwersytet Rolniczy  
ul. Balicka 253c 30-198 Kraków  
e-mail: rmbitner@cyf-kr.edu.pl

Recenzent: Prof. dr hab. inż. Ryszard Hycner