

DARIUSZ R. MAŃKOWSKI**TADEUSZ OLEKSIK**

Pracownia Ekonomiki Nasiennictwa i Hodowli Roślin

Zakład Nasiennictwa i Nasionoznawstwa

Instytut Hodowli i Aklimatyzacji Roślin w Radzikowie

Czynniki determinujące stosowanie kwalifikowanego materiału siewnego w gospodarstwach rolnych

Factors determining application of certified seed material in farms

Użycie przez rolnika kwalifikowanego materiału siewnego może być uzależnione od wielu czynników; ekonomicznych (cena materiału siewnego, dochodowość i opłacalność produkcji), technologicznych (stosowanej technologii i poziomu kultury uprawy), społecznych (wykształcenie, wiedza, doświadczenie), itp. Dodatkowo czynniki te mogą być ze sobą skorelowane. Trudno jest więc znaleźć jednoznaczną odpowiedź od czego zależy użycie kwalifikowanego materiału siewnego przez rolnika. W badaniach podjęto próbę identyfikacji czynników technologicznych wpływających na wykorzystanie kwalifikowanego materiału siewnego w uprawie głównych ziemiopłodów. Analizy wykonano w oparciu o dane pochodzące z badań ankietowych przeprowadzonych wśród rolników prowadzących rachunkowość rolniczą na potrzeby Instytutu Ekonomiki Rolnictwa i Gospodarki Żywnościowej. Badania prowadzono w latach 1992–2003. Łącznie, w tym okresie, obejmowały one 38 500 pól uprawnych. Uwzględniono dane dotyczące powierzchni całego gospodarstwa, powierzchni gruntów ornych w gospodarstwie, powierzchni uprawy gatunku, odległości gospodarstwa od punktu zaopatrzenia w materiał siewny, kwasowości gleby, jakości stanowiska, dawek nawożenia mineralnego, nawożenia organicznego, ilości wysiewu, terminu siewu oraz liczby zabiegów ochrony chemicznej. Do identyfikacji czynników determinujących wykorzystanie kwalifikowanego materiału siewnego wykorzystano analizę regresji logistycznej (PROC LOGISTIC) wykonaną w Systemie SAS® w wersji 9.1. Opracowano odrębne modele dla każdego z głównych ziemiopłodów (pszenicy ozimej, pszenicy jarej, jęczmienia jarego, pszenżyta ozimego oraz ziemniaków). Dobór zmiennych do modeli wykonano z wykorzystaniem kryteriów informacyjnych AIC (Akaike Information Criterion) oraz SBC (Schwarz Bayesian Criterion). Na podstawie przeprowadzonej analizy regresji logistycznej oceniono wpływ każdego z badanych czynników na wykorzystanie kwalifikowanego materiału siewnego w gospodarstwie.

Słowa kluczowe: badania ankietowe, regresja logistyczna, materiał siewny, jęczmień, pszenica, pszenżyto, ziemniaki

The farmers' decisions on the use of certified seed are influenced by numerous factors, e.g. economical (price of sowing material, farm profit), technological (soil and crop management systems), social (education, knowledge, experience) etc. The mentioned factors are usually correlated

(autocorrelation), which disturbs to give a simple answer to what the use of certified seed depends on. An attempt has been done at identification of some technological factors, which are significant for the main crops growers. The analysis based on survey research of the years 1992–2003, which was conducted on 38 500 fields of those farmers, who kept accountancy to meet the needs of Agriculture and Food Economy Institute. The data included: whole farm area, arable area, area under a specified crop, distance of a farm to a place of seed supply, soil acidity, soil quality, organic manuring, NPK fertilization, use of pesticides, rate and time of sowing. The data were subjected to the analysis of logistic regression (PROC LOGISTIC of the SAS[®] version 9.1) in order to determine the factors relevant for the use of certified seed material. The separate models were suggested for each of the studied crops (winter and spring wheat, spring barley, winter triticale, potato). Selection of variables for each of the models was performed using the AIC (Akaike Information Criterion) and SBC (Schwarz Bayesian Criterion). The use of certified seed was most closely related to the input factors, in particular to the mineral fertilization level and the number of fungicide treatments. The link was stated also with the production scale and total area of a farm.

Key words: barley, logistic regression, potatoes, seed material, survey research, triticale, wheat

WSTĘP

Materiał siewny uważany jest za ogniwo transmisji postępu odmianowego do produkcji polowej (Krzymuski, 1991, 1993, 1998; Oleksiak, 1998, 2000). Często jedną z barier w powszechnym wykorzystaniu kwalifikowanego materiału siewnego jest jego cena. Dzieje się tak, mimo że na podstawie wcześniejszych analiz (Oleksiak, 1998) podobnie jak i badań Instytutu Ekonomiki Rolnictwa i Gospodarki Żywnościowej, można stwierdzić, że koszt materiału siewnego zbóż stanowiący około 10–12% bezpośrednich kosztów produkcji jest względnie stałą i niewielką pozycją w relacji do kosztów nawozów czy paliwa (Oleksiak, 1998, Skarżyńska i in., 1996). Kwalifikowany materiał siewny łatwo jednak zastąpić własnym ziarnem. Między innymi, dlatego rolnicy często wykorzystują do siewu materiał własny o zdecydowanie gorszej jakości, lecz tańszy, bądź też kupują materiał niekwalifikowany.

Dla firm nasiennych duże znaczenie ma więc utrzymanie zadowolającego udziału w krajowym rynku nasion. Im ten udział będzie wyższy, tym wyższe będą dochody firm. Konkurencja dla krajowych firm nasiennych ze strony firm zagranicznych jest jednak bardzo groźna. W roku 2003 udział największych zagranicznych firm nasiennych w krajowym rynku materiału siewnego roślin zbożowych wynosił około 4%, w roku 2004 — około 11%, a w roku 2005 już około 13% (Oleksiak, 2004, 2006). Kluczowym narzędziem mającym na celu pomoc w utrzymaniu pozycji krajowych firm nasiennych na rynku staje się marketing. Jest to jednak działalność kosztowna. Potrzeba przemyślanych strategii i ukierunkowanych akcji promocyjnych. Niestety, brak jest kompleksowego rozpoznania rynku potencjalnych odbiorców, a kierowanie akcji marketingowych do całego rynku, może okazać się, przy skromnym budżecie firm krajowych, niemożliwe do wykonania. Istnieje więc potrzeba rozpoznania rynku i odnalezienia takich cech gospodarstw rolniczych, poza ceną materiału siewnego, które mogłyby wskazywać na to czy dany rolnik byłby zainteresowany kupnem kwalifikowanego materiału siewnego, czy też nie. Badania mające na celu rozpoznanie sytuacji rynkowej najczęściej prowadzi się za

pomocą szeroko zakrojonych badań ankietowych, prowadzonych wśród uczestników rynku.

W badaniach ankietowych często pozyskuje się dane o rozkładzie dwumianowym, najczęściej binarnym (0, 1), oznaczające wystąpienie, bądź niewystąpienie jakiegoś faktu w danym zjawisku. Jeżeli chcemy ustalić jak, pod wpływem zmiennych przyczynowych kształtuje się zmienna zależna o charakterze ciągłym, stosujemy zwykle analizę liniowej regresji wielokrotnej. Jednak tej metody nie możemy stosować w przypadku, gdy zmienna zależna posiada rozkład binarny. Wówczas możemy stosować regresję logistyczną. Metoda ta jest efektywnie stosowana w psychologii, medycynie oraz ekonomii, natomiast nie jest zbyt rozpowszechniona w badaniach rolniczych.

MATERIAŁ I METODY

Materiał do analiz stanowiły wyniki badań ankietowych gospodarstw indywidualnych z lat 1992–2003. Do analiz wykorzystano dane dotyczące 38 500 pól uprawnych, na których uprawiano główne ziemniopłydy: pszenicę ozimą, pszenicę jara, jęczmień jary, pszenżyto ozime i ziemniaki. Tabela 1 zawiera zestawienie wykorzystanych w analizach zmiennych z badań ankietowych. Są to zmienne mówiące o gospodarstwie i polu uprawnym (powierzchnia pola uprawnego, powierzchnia gospodarstwa, powierzchnia gruntów ornych w gospodarstwie, odległość gospodarstwa od punktu zaopatrzenia w materiał siewny, odczyn gleby oraz jakość stanowiska / gleby), mówiące o nakładach na produkcję (liczba zabiegów pestycydami — fungicydami, herbicydami i insektycydami, nawożenie organiczne, nawożenie mineralne NPK) oraz zmienne dotyczące materiału siewnego (jakość materiału siewnego, ilość wysiewu, termin siewu).

Do opisu wpływu grupy cech (tzw. zmiennych przyczynowych) na zachowanie się jednej cechy (tzw. zmiennej zależnej) najczęściej stosuje się analizę funkcji regresji wielokrotnej. Metoda ta jednak wymaga spełnienia założenia, iż zmienna zależna ma rozkład ciągły i zbliżony do rozkładu normalnego. Wówczas funkcja regresji przyjmuje postać:

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots \quad (1)$$

gdzie:

Y — zmienna zależna;

X_1, X_2, \dots — zmienne przyczynowe;

β_0 — stała regresji;

β_1, β_2 — cząstkowe współczynniki regresji.

Dla innych rozkładów zmiennej zależnej można zastosować różne modele zależnie od tego rozkładu.

W przypadku rozkładu binarnego, zmienna Y może przyjmować tylko dwie wartości 0 lub 1. Zmienną zależną jest rodzaj (jakość) materiału siewnego. Zmienna ta przyjmuje wartości: 1 — gdy stosowany był materiał kwalifikowany oraz 0 — gdy na pole wysiano materiał niekwalifikowany własny bądź z wymiany lub zakupu. Funkcja zmiennej binarnej przyjmuje postać (Agresti, 2002):

$$Y = \frac{e^{\beta_0 + \beta \cdot X}}{1 + e^{\beta_0 + \beta \cdot X}} \quad (2)$$

gdzie:

Y — zmienna zależna;

X — zmienna przyczynowa;

β_0 — stała regresji,

β — wektor parametrów.

Przekształcając funkcję regresji logistycznej (wzór 1), możemy wyznaczyć wartości szans (ang. odds) jako:

$$\frac{Y}{1-Y} = e^{\beta_0 + \beta \cdot X} \quad (3)$$

a następnie jako:

$$\ln\left(\frac{Y}{1-Y}\right) = \beta_0 + \beta \cdot X \quad (4)$$

Ten zapis nosi nazwę logitu:

$$\text{logit}[Y] = \ln\left(\frac{Y}{1-Y}\right) = \beta_0 + \beta \cdot X \quad (5)$$

Logit jest naturalnym parametrem rozkładu dwumianowego. Jeżeli wartość zmiennej Y musi wynosić 0 lub 1, to logit może przyjmować dowolną wartość ze zbioru liczb rzeczywistych.

W przypadku regresji logistycznej wielokrotnej funkcja przyjmuje postać (Der i Everitt, 2002):

$$\text{logit}[Y] = \ln\left(\frac{Y}{1-Y}\right) = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots \quad (6)$$

gdzie:

Y — zmienna zależna;

X_1, X_2, \dots — zmienne przyczynowe;

β_0 — stała regresji;

β_1, β_2, \dots — cząstkowe współczynniki regresji.

Logit prawdopodobieństwa jest logarytmem naturalnym szansy (ang. odd) zajścia danego zdarzenia (Der i Everitt, 2002). Dla $\boldsymbol{\beta}' = [\beta_0, \beta_1, \beta_2, \dots, \beta_k]$ oraz wektora parametrów dla i -tych wartości jako $\mathbf{X}'_i = [1, X_{i1}, X_{i2}, \dots, X_{ik}]$, przewidywane prawdopodobieństwo w postaci funkcji liniowej można zapisać wzorem:

$$Y(\boldsymbol{\beta}'\mathbf{X}_i) = \frac{e^{(\boldsymbol{\beta}'\mathbf{X}_i)}}{1 + e^{(\boldsymbol{\beta}'\mathbf{X}_i)}} \quad (7)$$

Zmienne wykorzystane w analizie regresji logistycznej
Variables used in the logistic regression analysis

Nazwa zmiennej Variable name	Opis Description	Ocena Evaluation
AREA	Powierzchnia pola, na którym uprawia się dany gatunek Area of field, on which the species was cultivated	ha
DISTANCE	Odległość gospodarstwa od punktu zaopatrzenia w materiał siewny Distance from farm to seed store	km
F_AREA	Powierzchnia gospodarstwa Farm area	ha
FUNGI	Liczba zabiegów fungicydami Number of fungicides operations	krotność zabiegów w okresie wegetacji number of operations in vegetation period
HERB	Liczba zabiegów herbicydami Number of herbicides operations	krotność zabiegów w okresie wegetacji number of operations in vegetation period
INSECT	Liczba zabiegów insektydami Number of pesticides operations	krotność zabiegów w okresie wegetacji number of operations in vegetation period
K	Dawka nawożenia potasem w czystym składniku K fertilization dose, pure element	kg·ha ⁻¹
MANURE	Nawożenie obornikiem Organic manure	liczba lat wstecz od momentu ostatniego nawożenia obornikiem (0 — w danym roku; 1 — w roku poprzedzającym, itd.) number of years after the last organic manuring (0 — the same year, 1 — one year after, etc.)
N	Dawka nawożenia azotem w czystym składniku N fertilization dose, pure element	kg·ha ⁻¹
NPK	Łączna dawka nawożenia azotem, fosforem i potasem w czystym składniku N+P+K fertilization dose, pure element	kg·ha ⁻¹
P	Dawka nawożenia fosforem w czystym składniku P fertilization dose, pure element	kg·ha ⁻¹
PEST	Łączna liczba zabiegów chemicznej ochrony roślin Summarized number of chemical control operations	krotność zabiegów w okresie wegetacji number of operations in vegetation period
SEED_NO	Ilość wysiewu Sowing rate	kg·ha ⁻¹
SEED_QUALITY	Użyty materiał siewny Quality of seed material	0 — materiał niekwalifikowany, 1 — materiał kwalifikowany; 0 — non-qualified seed material, 1 — qualified seed material
SOIL_AC	Kwasowość gleby Soil pH	skala (1 — gleba bardzo kwaśna, 2 — gleba kwaśna, 3 — gleba lekko kwaśna, 4 — gleba obojętna, 5 — gleba zasadowa) scale (1 — very low pH, 2 — low pH, 3 — moderately low pH, 4 — neutral pH)
SOIL_P	Jakość gleby wyrażona w punktach waloryzacji rolniczej przestrzeni produkcyjnej (Witek, 1981) Soil quality in grades of agricultural production area evaluation	punkty (0 – 100) points (0 – 100)
SOWING_D	Termin siewu w dniach Sowing date in days	liczba dni od 1 stycznia number of days after 1 January
U_AREA	Powierzchnia gruntów ornych w gospodarstwie Farm arable area	ha

Podczas gdy logit może przyjmować dowolne wartości rzeczywiste, prawdopodobieństwo może przyjmować tylko wartości z zakresu $0 \leq Y(\beta'X_i) \leq 1$. W modelu regresji logistycznej parametr β_1 powiązany ze zmienną przyczynową X_i w postaci $e^{\beta_1 X_i}$ jest szansą (ang. odd) na to, że $Y = 1$, gdy X_i wzrośnie o jeden (Der i Everitt, 2002).

Do oszacowania parametrów modelu zapisanego wzorem 6 wykorzystuje się metodę maksymalnej wiarygodności (ang. maximum likelihood). Funkcja wiarygodności dla regresji logistycznej ma postać:

$$l(\beta, Y) = \sum_i Y_i \cdot \ln[Y(\beta'X_i)] + (1 - Y_i) \cdot \ln[1 - Y(\beta'X_i)] \quad (8)$$

gdzie $Y' = [Y_1, Y_2, \dots, Y_n]$ jest obserwowanymi n wartościami dychotomicznej (binarnej) zmiennej zależnej. Funkcja wiarygodności dla regresji logistycznej jest maksymalizowana numerycznie z użyciem algorytmu iteracyjnego (Der i Everitt, 2002).

Budowa optymalnego modelu regresyjnego polega na wyborze jak najmniejszej liczby zmiennych przyczynowych przy zachowaniu jak najwyższej zdolności opisującej modelu. W statystyce matematycznej istnieje szereg różnych miar pozwalających na jak najlepszy, pod różnym względem, dobór zmiennych do modelu regresyjnego. Niestety większość z nich opracowana jest z myślą o analizie funkcji regresji wielokrotnej. W przypadku analizy regresji logistycznej możemy zastosować tak zwane kryteria informacyjne: *AIC* oraz *SBC* (Rawlings i in., 2001).

Kryterium informacyjne Akaike (1969) (*AIC* — ang. Akaike Information Criterion) wyznacza się jako:

$$AIC(p') = n \cdot \ln(SSE_p) + 2 \cdot p' - n \cdot \ln(n) \quad (9)$$

gdzie:

$AIC(p')$ — wartość *AIC* dla modelu z p' zmiennymi przyczynowymi w modelu, wliczając stałą;

n — liczba obserwacji;

SSE_p — suma kwadratów odchyłeń błędu dla modelu zawierającego wszystkie dostępne zmienne przyczynowe i stałą;

p' — liczba zmiennych w modelu wliczając stałą.

Wartość *AIC* wyznacza się dla modeli z kombinacją wszystkich zmiennych w modelu (zaczynając od modelu tylko ze stałą do modelu, w skład którego wchodzi wszystkie zmienne przyczynowe i stała), następnie za model optymalny uznaje się ten dla którego wyznaczona wartość *AIC* jest najmniejsza.

Kryterium informacyjne *SBC* (ang. Schwarz Bayesian Criterion), opracowane przez Schwarza (1978) oblicza się jako:

$$SBC(p') = n \cdot \ln(SSE_p) + [\ln(n)] \cdot p' - n \cdot \ln(n) \quad (10)$$

gdzie:

$SBC(p')$ — wartość *SBC* dla modelu z p' zmiennymi przyczynowymi w modelu, wliczając stałą;

n — liczba obserwacji;

SSE_p — suma kwadratów odchylenia błędu dla modelu zawierającego wszystkie dostępne zmienne przyczynowe i stałą;

p' — liczba zmiennych w modelu wliczając stałą.

SBC podobnie jak AIC jest wyznaczane dla modeli wszystkich kombinacji zmiennych przyczynowych. Model, który ma najniższą wartość SBC jest uznawany za model optymalny.

Obliczenia wykonano z wykorzystaniem Systemu SAS® w wersji 9.1 (Allison, 1999; Stokes i Davis, 2000; Der i Everitt, 2002; SAS Institute Inc., 2004 a, 2004 b, 2004 c; Cody i Sumith, 2005).

Procedura LOGISTIC w Systemie SAS®

Do obliczeń wykorzystano procedurę LOGISTIC, która jest składową modułu SAS®STAT (SAS Institute Inc., 2004 c). Jej składnia w najprostszej formie jest następująca:

```
PROC LOGISTIC <OPCJE>;
  MODEL zmienna_zalezna = zmienne_przyczynowe /<OPCJE>;
RUN;
```

Ponieważ procedura ta wykonuje obliczenia dla jednego z góry określonego modelu i nie pozwala na wybór najlepszego modelu pod względem kryteriów *SBC* i *AIC* wartości tych kryteriów wyznaczono w oddzielnych podprogramach zwanych 'data-stepami'. Poniżej znajduje się przykładowy kod programu, dla pszenicy ozimej, w wyniku którego dla wszystkich możliwych modeli zostaną wyznaczone wartości kryteriów *SBC* i *AIC*:

```
ods listing close;
ods output bestsubsets = modele;
proc logistic data = pszenica_ozima desc;
  model seed_quality = f_area u_area area manure n p k
    seed_no herb fungi insect soil_p npk sowing_d
    est distance soil_ac / selection = score;
run;
ods output close;
ods listing;
data _null_;
  set pszenica_ozima nobs = il;
  call symput ('obs', put(il, best12.-L));
  stop;
run;
data modele;
  set modele;
  sbc = -scorechisq + log(&obs) * (numberofvariables + 1);
  aic = -scorechisq + 2 * (numberofvariables + 1);
run;
```

Następnie na podstawie wyznaczonych wartości wybrano modele o najmniejszych wartościach *SBC* i *AIC*. Dla tych modeli wykonano szczegółową analizę regresji logistycznej. Poniżej znajduje się przykładowy kod programu wykonującego taką analizę, dla pszenicy ozimej, dla modelu ze stałą oraz modelu bez stałej regresji:

```
proc logistic data = pszenica_ozima desc;
  model seed_quality = k herb fungi;
run;
proc logistic data = pszenica_ozima desc;
  model seed_quality = k herb fungi / noint;
run;
```

W przeprowadzonych analizach wzięto pod uwagę dwa uzyskane modele dla każdego z analizowanych gatunków roślin. Jeden model uzyskano przy zastosowaniu kryterium *AIC*, drugi z kolei przy zastosowaniu kryterium *SBC*. Dla każdego z kryteriów wyznaczono parametry funkcji regresji logistycznej z uwzględnieniem stałej regresji w modelu oraz bez stałej w modelu. Biorąc pod uwagę statystyki wyznaczone dla stałej regresji wybrano bądź wariant ze stałą, bądź bez stałej w modelu.

OMÓWIENIE WYNIKÓW

Pszenica ozima

Przy zastosowaniu kryterium *SBC* najlepszy model zawierał stałą i trzy zmienne (*SBC* = -219,698). W modelu znalazły się zmienne: nawożenie potasem, liczba zabiegów ochrony chemicznej herbicydami oraz fungicydami.

Tabela 2

Pszenica ozima — wyniki analizy ocen maksymalnej wiarygodności oraz ocen ilorazu szans
Winter wheat — results of analysis of maximum likelihood estimates and odds ratio estimates

Parametr Parameter	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation
	model ze stałą — model with intercept			model bez stałej — model without intercept		
Model z trzema zmiennymi Model with three variables						
Stała - Intercept	-2,1527	107,4145**	—	—	—	—
K	0,0123	35,8046**	1,012	0,00304	2,9634	1,007
HERB	0,6395	19,6132**	1,896	-0,6099	55,7650**	0,638
FUNGI	0,9768	107,8146**	2,656	0,9854	110,2795**	3,220
Model z pięcioma zmiennymi Model with five variables						
Stała - Intercept	-2,1235	100,1427**	—	—	—	—
F_AREA	-0,0763	7,1778**	0,926	-0,1246	17,8778**	0,883
U_AREA	0,0846	8,1171**	1,088	0,1319	18,5513**	1,141
K	0,0118	31,9883**	1,012	0,00308	2,9061	1,003
HERB	0,5904	16,6142**	1,805	-0,5966	49,5325**	0,551
FUNGI	0,9202	90,4265**	2,510	0,9327	92,0901**	2,541

** Istotne przy $\alpha = 0,01$; ** Significant at $\alpha = 0.01$

A najlepszy model przy zastosowaniu kryterium *AIC* (-246,307) był zbudowany ze stałej i pięciu zmiennych: powierzchnia gospodarstwa, powierzchnia gruntów ornych, nawożenie potasem, liczba zabiegów ochrony chemicznej herbicydami oraz fungicydami. Wyniki analizy regresji logistycznej pozwoliły na odrzucenie hipotezy globalnej o zerowych wartościach wszystkich współczynników β (chi-kwadrat ilorazu wiarygodności dla modelu z trzema zmiennymi i stałą wynosił 168,28**, a dla modelu z pięcioma zmiennymi i stałą wynosił 295,6426**). Szczegółowe oceny wartości parametrów modelu regresji logistycznej przedstawiono w tabeli 2.

Pszonica jara

Model zbudowany w oparciu o kryterium *SBC* (-61,0341) zawierał cztery zmienne (stała była nieistotna statystycznie), były to: powierzchnia gruntów ornych, liczba zabiegów fungicydami, jakość gleby w punktach, termin siewu. Przy zastosowaniu kryterium *AIC* (-93,8335) zbudowano model zawierający osiem zmiennych (stała również była nieistotna statystycznie): powierzchnia gospodarstwa, powierzchnia gruntów ornych, nawożenie fosforem, ilość wysiewu, liczba zabiegów fungicydami i insektycydami, jakość gleby w punktach oraz termin siewu.

Tabela 3

Pszonica jara — wyniki analizy ocen maksymalnej wiarygodności oraz ocen ilorazu szans Spring wheat — results of analysis of maximum likelihood estimates and odds ratio estimates

Parametr Parameter	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation
	model zawierający stałą model with intercept			model nie zawierający stałej model without intercept		
Model z czterema zmiennymi Model with four variables						
Stała - Intercept	0,1006	0,0155	—	—	—	—
U_AREA	0,00728	6,2009*	1,007	0,00730	6,2441*	1,013
FUNGI	0,7806	30,9420**	2,183	0,7838	32,2773**	2,870
SOIL_P	0,0234	25,7154**	1,024	0,0237	32,8200**	1,032
SOWING_D	-0,0258	12,6799**	0,974	-0,0250	75,9902**	0,981
Model z ośmioma zmiennymi Model with eight variables						
Stała - Intercept	1,2208	1,5311	—	—	—	—
F_AREA	-0,1181	5,1271*	0,889	-0,1150	4,8558*	0,891
U_AREA	0,1258	5,5450*	1,134	0,1230	5,2941*	1,131
P	0,00858	7,3671**	1,009	0,00902	8,2614**	1,009
SEED_NO	-0,00469	6,0531*	0,995	-0,00344	4,5435*	0,997
FUNGI	0,6560	17,6186**	1,927	0,6641	18,1176**	1,943
INSECT	0,4321	1,9937	1,541	0,4009	1,7482	1,493
SOIL_P	0,0220	21,2746**	1,022	0,0241	29,2125**	1,024
SOWING_D	-0,0255	11,7598**	0,975	-0,0181	17,3332**	0,982

* Istotne przy $\alpha = 0,05$; * Significant at $\alpha = 0.05$

** Istotne przy $\alpha = 0,01$; ** Significant at $\alpha = 0.01$

Wyniki analizy regresji logistycznej pozwoliły na odrzucenie hipotezy globalnej o zerowych wartościach wszystkich współczynników β (chi-kwadrat ilorazu wiarygodności

dla modelu z czterema zmiennymi bez stałej wynosił 156,6871**, a dla modelu z ośmioma zmiennymi bez stałej wynosił 174,6902**). Szczegółowe oceny wartości parametrów modelu regresji logistycznej przedstawiono w tabeli 3.

Jęczmień jary

Najlepszy model zbudowany w oparciu o kryterium *SBC* (-44,1257) zawierał pięć zmiennych bez stałej. Do modelu weszły: powierzchnia gruntów ornych, nawożenie potasem, ilość wysiewu, liczba zabiegów fungicydami oraz jakość gleby. Przy zastosowaniu kryterium *AIC* (-79,6014) do modelu weszło dziewięć zmiennych bez stałej, były to: powierzchnia gospodarstwa, powierzchnia gruntów ornych, liczba lat od nawożenia obornikiem, nawożenie azotem i potasem, ilość wysiewu, liczba zabiegów fungicydami, jakość gleby oraz termin siewu. Wyniki analizy regresji logistycznej pozwoliły na odrzucenie hipotezy globalnej o zerowych wartościach wszystkich współczynników β (chi-kwadrat ilorazu wiarygodności dla modelu z pięcioma zmiennymi bez stałej wynosił 180,8222**, a dla modelu z dziewięcioma zmiennymi bez stałej wynosił 196,25**). Szczegółowe oceny wartości parametrów modelu regresji logistycznej przedstawiono w tabeli 4.

Tabela 4

Jęczmień jary — wyniki analizy ocen maksymalnej wiarygodności oraz ocen ilorazu szans Spring barley — results of analysis of maximum likelihood estimates and odds ratio estimates

Parametr Parameter	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation
	model zawierający stałą model with intercept			model nie zawierający stałej model without intercept		
Model z pięcioma zmiennymi Model with five variables						
Stała - Intercept	-0,3871	0,4681	—	—	—	—
U_AREA	0,00795	9,0105**	1,008	0,00764	8,5719**	1,008
K	0,00864	10,8640**	1,009	0,00813	10,5272**	1,008
SEED_NO	-0,00898	13,1191**	0,991	-0,0105	70,3356**	0,990
FUNGI	0,5112	11,3825**	1,667	0,5051	11,1550**	1,657
SOIL_P	0,0136	11,0316**	1,014	0,0124	11,1325**	1,013
Model z dziewięcioma zmiennymi Model with nine variables						
Stała - Intercept	1,2681	1,4642	—	—	—	—
F_AREA	-0,0757	4,4750*	0,927	-0,0725	4,1324*	0,930
U_AREA	0,0878	5,5763*	1,092	0,0844	5,1977*	1,088
MANURE	-0,1126	4,6144*	0,893	-0,0999	3,8029	0,905
N	0,00482	3,3451	1,005	0,00548	4,5253*	1,005
K	0,00768	7,8869**	1,008	0,00814	8,9864**	1,008
SEED_NO	-0,00889	11,9812**	0,991	-0,00770	10,6603**	0,992
FUNGI	0,4765	8,8110**	1,610	0,4829	9,0928**	1,621
SOIL_P	0,0131	9,2025**	1,013	0,0148	13,0886**	1,015
SOWING_D	-0,0159	3,5875	0,984	-0,00754	2,5088	0,992

* Istotne przy $\alpha = 0,05$; * Significant at $\alpha = 0.05$

** Istotne przy $\alpha = 0,01$; ** Significant at $\alpha = 0.01$

Pszenżyto ozime

Model zbudowany w oparciu o kryterium SBC (-11,6255) zawierał stałą dwie zmienne nawożenie potasem oraz liczba zabiegów fungicydami. Przy zastosowaniu kryterium AIC (-26,4126) zbudowano model zawierający stałą i pięć zmiennych: nawożenie potasem, liczba zabiegów fungicydami i insektycydami, jakość gleby oraz odległość od punktu zaopatrzenia w kwalifikowany materiał siewny. Wyniki analizy regresji logistycznej pozwoliły na odrzucenie hipotezy globalnej o zerowych wartościach wszystkich współczynników β (chi-kwadrat ilorazu wiarygodności dla modelu z dwoma zmiennymi i stałą wynosił 31,8016**, a dla modelu z pięcioma zmiennymi i stałą wynosił 40,1755**). Szczegółowe oceny wartości parametrów modelu regresji logistycznej przedstawiono w tabeli 5.

Tabela 5

Pszenżyto ozime — wyniki analizy ocen maksymalnej wiarygodności oraz ocen ilorazu szans
Winter triticale — results of analysis of maximum likelihood estimates and odds ratio estimates

Parametr Parameter	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation
	model zawierający stałą model with intercept			model nie zawierający stałej model without intercept		
Model z dwoma zmiennymi Model with two variables						
Stała - Intercept	-1,2100	55,6453**	—	—	—	—
K	0,0127	15,8980**	1,013	-0,00499	5,6659*	0,995
FUNGI	0,5567	8,1123**	1,745	0,3140	2,7853	1,369
Model z pięcioma zmiennymi Model with five variables						
Stała - Intercept	-1,9382	23,1851**	—	—	—	—
K	0,0124	14,9232**	1,012	0,00789	6,8839**	1,008
FUNGI	0,6817	9,9885**	1,977	0,6917	10,4701**	1,997
INSECT	-0,8491	3,1804	0,428	-0,9541	4,1269*	0,385
SOIL_P	0,00995	2,5729	1,010	-0,0155	20,0811**	0,985
DISTANCE	0,0192	2,7143	1,019	-0,00937	0,8658	0,991

* Istotne przy $\alpha = 0,05$; * Significant at $\alpha = 0.05$

** Istotne przy $\alpha = 0,01$; ** Significant at $\alpha = 0.01$

Ziemniaki

Najlepszy model regresji logistycznej pod względem kryterium SBC (-215,560) zawierał siedem zmiennych, były to: powierzchnia gruntów ornych, nawożenie azotem, ilość wysiewu, liczba zabiegów herbicydami i fungicydami, jakość gleby oraz termin sadzenia.

Natomiast najlepszy pod względem kryterium AIC (-263,023) zawierał dziewięć zmiennych, były to: powierzchnia gruntów ornych, powierzchnia pola, nawożenie azotem, ilość wysiewu, liczba zabiegów herbicydami i fungicydami, jakość gleby, termin sadzenia oraz odczyn gleby. Wyniki analizy regresji logistycznej pozwoliły na odrzucenie hipotezy globalnej o zerowych wartościach wszystkich współczynników β (chi-kwadrat ilorazu wiarygodności dla modelu bez stałej wynosił 951,6936**, a dla modelu bez stałej wynosił

957,7404**). Szczegółowe oceny wartości parametrów modelu regresji logistycznej przedstawiono w tabeli 6.

Tabela 6

Ziemniaki — wyniki analizy ocen maksymalnej wiarygodności oraz ocen ilorazu szans
Potatoes — results of analysis of maximum likelihood estimates and odds ratio estimates

Parametr Parameter	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation	Ocena Estimate	Chi-kwadrat Walda Wald chi-square	Ocena ilorazu szans Odds ratio estimation
	model zawierający stałą model with intercept			model nie zawierający stałej model without intercept		
Model z siedmioma zmiennymi Model with seven variables						
Stała - Intercept	-0,8892	1,1934	—	—	—	—
U_AREA	0,00681	9,6216**	1,007	0,00649	8,8776**	1,007
N	-0,00508	15,7445**	0,995	-0,00527	17,2159**	0,955
SEED_NO	0,000713	15,0683**	1,001	0,000611	14,9680**	1,001
HERB	0,4552	14,1072**	1,576	0,4478	13,6768**	1,565
FUNGI	0,4842	64,3103**	1,623	0,4858	64,7110**	1,625
SOIL_P	0,0161	33,7310**	1,016	0,0154	32,5640**	1,016
SOWING_D	-0,0311	29,0755**	0,969	-0,0362	120,0120**	0,964
Model z dziewięcioma zmiennymi Model with nine variables						
Stała - Intercept	-1,5598	3,3206	—	—	—	—
U_AREA	0,00563	5,9934*	1,006	0,00521	5,1864*	1,010
AREA	0,0914	4,1834*	1,096	0,0874	4,0465*	1,230
N	-0,00547	17,5467**	0,995	-0,00574	19,5629**	0,997
SEED_NO	0,000759	16,8233**	1,001	0,000585	13,6819**	1,001
HERB	0,4235	12,0304**	1,527	0,4173	11,6518**	1,929
FUNGI	0,4729	60,1492**	1,605	0,4782	61,6234**	1,818
SOIL_P	0,0156	31,6332**	1,016	0,0146	28,8050**	1,020
SOWING_D	-0,0303	27,6568**	0,970	-0,0386	119,1347**	0,969
SOIL_AC	0,1530	5,8272*	1,165	0,1176	4,8111*	1,266

* Istotne przy $\alpha = 0,05$; * Significant at $\alpha = 0.05$

** Istotne przy $\alpha = 0,01$; ** Significant at $\alpha = 0.01$

DYSKUSJA

Najsilniejszy związek ze stosowaniem kwalifikowanego materiału siewnego pszenicy ozimej stwierdzono dla takich czynników jak: nawożenie potasem, ilość zabiegów ochrony chemicznej herbicydami i fungicydami, w dalszej kolejności zidentyfikowano wpływ wielkości gospodarstwa oraz powierzchni gruntów ornych w gospodarstwie. Wszystkie te cechy, za wyjątkiem powierzchni gospodarstwa, mają dodatni wpływ na szansę użycia przez rolnika kwalifikowanego materiału siewnego, czyli im te cechy przyjmują wyższe wartości, tym bardziej wzrasta szansa na to, że rolnik zakupi kwalifikowany materiał siewny. Prawdopodobieństwo zastosowania materiału kwalifikowanego zwiększa się wraz ze wzrostem liczby zabiegów fungicydami, herbicydami, wraz ze wzrostem poziomu nawożenia potasem i zwiększaniem powierzchni gruntów ornych w gospodarstwie. Kierując akcją promocyjną kwalifikowanego materiału siewnego odmian pszenicy ozimej

firmy nasienne powinny więc kierować do rolników, którzy w prowadzonej przez siebie produkcji polowej wykorzystują fungicydy, herbicydy oraz znaczne dawki nawożenia potasem jak również do gospodarstw o większej skali produkcji. Ujemna korelacja między stosowaniem kwalifikowanego materiału siewnego a powierzchnią gospodarstwa wynikała z dużego udziału trwałych użytków zielonych w powierzchni tych gospodarstw.

W przypadku uprawy pszenicy jarej istotny związek ze stosowaniem kwalifikowanego materiału siewnego stwierdzono dla powierzchni gruntów ornych w gospodarstwie, liczby zabiegów fungicydami, jakości stanowiska oraz terminu siewu, wpływ miały również powierzchnia gospodarstwa, nawożenie fosforem, ilość wysiewu oraz ilość zabiegów insektycydami, przy czym ta ostatnia cecha nie wykazała istotnej wartości testu χ^2 — Walda. Dodatni wpływ na szansę wykorzystania kwalifikowanego materiału siewnego miały wszystkie wymienione cechy za wyjątkiem terminu siewu (im termin późniejszy, tym mniejsza szansa), ilości wysiewu i powierzchni gospodarstwa. Najsilniejszy wpływ miały: ilość zabiegów fungicydami oraz ilość zabiegów insektycydami. Kierując akcją promocyjną do rolników uprawiających pszenicę jarą firmy nasienne powinny zwrócić uwagę na rolników stosujących intensywne technologie, przy czym ważne jest by materiał siewny nie był wprowadzony na rynek zbyt późno, gdyż rolnicy stosujący siew w późniejszym terminie rzadziej sięgają po materiał kwalifikowany.

Najsilniejszy związek ze stosowaniem kwalifikowanego materiału siewnego w uprawie jęczmienia jarego wykazywały: liczba zabiegów fungicydami, jakość gleby, powierzchnia gruntów ornych, nawożenie potasem oraz nawożenie azotem. Nawożenie azotem i termin siewu nie wykazały jednak istotnej statystyki testowej. Wpływ ujemny wykazały: ilość wysiewu, powierzchnia gospodarstwa, nawożenie organiczne obornikiem, termin siewu. Pozostałe cechy istotnie wpływając na użycie kwalifikowanego materiału siewnego wykazywały wpływ dodatni. Przygotowując reklamy i promocje skierowane do producentów jęczmienia jarego, firmy nasienne powinny więc zwrócić szczególną uwagę na rolników, którzy stosują intensywne nawożenie potasem i ochronę chemiczną przed chorobami grzybowymi, mają gospodarstwa o dużym udziale gruntów ornych i dobrej jakości gleby.

W przypadku rolników uprawiających pszenżyto ozime, fakt użycia materiału kwalifikowanego istotnie wiązał się z nawożeniem potasem oraz liczbą zabiegów ochrony chemicznej przed chorobami grzybowym, znacznie mniejsze znaczenie, bo nie stwierdzono dla nich istotnych wartości statystyki testowej testu χ^2 — Walda, miały takie czynniki jak jakość gleby, odległość od gospodarstw do punktu zaopatrzenia w materiał siewny oraz ilość zabiegów ochrony chemicznej insektycydami. Przygotowując akcję promocyjną, firmy nasienne powinny więc kierować ją do rolników stosujących wysokie nawożenie, szczególnie potasem, oraz stosujących pełną ochronę chemiczną przed chorobami grzybowymi.

W przypadku ziemniaka największy wpływ na to czy rolnik zdecydował się na zakup kwalifikowanych sadzeniaków miały: powierzchnia gruntów ornych (wpływ dodatni), nawożenie azotem (wpływ ujemny), ilość wysiewu (dodatni), liczba zabiegów herbicydami (dodatni), liczba zabiegów fungicydami (dodatni), jakość gleby (dodatni), termin sadzenia (ujemny); w dalszym rzędzie wpływ miały: powierzchnia pola (dodatni) oraz odczyn gleby

(dodatni). Najsilniejszy wpływ wykazały: liczba zabiegów fungicydami oraz liczba zabiegów herbicydami. Świadczy to tym, iż rolnicy stosujący najwyższy poziom agrotechniki zdecydowanie częściej sięgają po materiał kwalifikowany.

Przedstawione analizy obejmują jedynie pewną część spośród krajowych gospodarstw rolnych. Są to gospodarstwa wielkości 20–30 ha, zajmujące się głównie produkcją towarową. Tak więc wyników przeprowadzonych analiz nie można przełożyć na wszystkie gospodarstwa rolne. Założeniem niniejszej pracy była prezentacja metody umożliwiającej firmom nasiennym sprecyzowanie i ukierunkowanie kampanii marketingowych na określonych odbiorców. Opracowanie to jest tylko próbą prezentacji tego typu analizy. Aby otrzymać wyniki lepiej przekładające się na ogół gospodarstw rolnych należałoby przeprowadzić badania ankietowe na szerszą skalę, z lepszym, bardziej przekrojowym doбором próby.

Na podstawie przedstawionych wyników można zaobserwować wyraźne prawidłowości, które powinny znaleźć odzwierciedlenie w przypadku ogółu gospodarstw rolniczych. Wyraźnie widać, że kwestia wykorzystania kwalifikowanego materiału siewnego najsilniej wiąże się z poziomem agrotechniki stosowanej w gospodarstwach. Co ciekawe na fakt zakupu kwalifikowanego materiału siewnego nie miała istotnego wpływu odległość od gospodarstwa do punktu zaopatrzenia w ten materiał. Uzyskane wyniki pozwalają na takie przygotowanie reklam i promocji, aby dotarły one do odbiorców, którzy zakupią kwalifikowany materiał siewny. Reklamy odmian i firm nasiennych można dla przykładu dołączać do środków ochrony roślin, w szczególności do fungicydów lub do nawozów.

WNIOSKI

1. Analizę zależności pomiędzy zmiennymi przyczynowymi a zmienną zależną o rozkładzie binarnym (zerojedynkowym, dychotomicznym) można przeprowadzić za pomocą analizy regresji logistycznej. Wybór właściwego modelu można znacznie ułatwić poprzez zastosowanie kryterium *AIC* lub *SBC*.
2. W analizowanych gatunkach roślin uprawnych fakt wykorzystywania przez rolników kwalifikowanego materiału siewnego najsilniej związany był z nakładowymi czynnikami produkcji, a przede wszystkim z nawożeniem mineralnym i liczbą zabiegów z wykorzystaniem fungicydów. Stwierdzono również związek ze skalą produkcji, im większa powierzchnia gruntów ornych w gospodarstwie, tym większe prawdopodobieństwo zastosowania kwalifikowanego materiału siewnego.
3. Zaproponowana metoda statystyczna może być pomocna w przygotowywaniu przez firmy nasienne, kampanii marketingowych nowych odmian i materiału siewnego.

LITERATURA

- Agresti A. 2002. Categorical data analysis. 2nd Edition. New Jersey, USA, John Wiley & Sons Inc.
Akaike H. 1969. Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21: 243 — 247.

- Allison P. D. 1999. Logistic Regression Using the SAS System — Theory and Application. New York, USA, SAS Publishing, SAS Institute Inc., John Wiley & Sons Inc.
- Cody R. P., Smith J. K. 2005. Applied Statistics and the SAS Programming Language. 5th Edition. Upper Saddle River, NJ, USA, Pearson Education Inc.
- Der G., Everitt B. S. 2002. A handbook of statistical analyses using SAS. 2nd Edition. USA, CRC Press LLC.
- Krzymuski J. 1991. Postęp odmianowy w produkcji zbóż w Polsce. Cz. III. Introdukcja z hodowli do produkcji. Biul. IHAR 177: 17 — 24.
- Krzymuski J. 1993. Ekonomiczne aspekty nasiennictwa. Biul. IHAR 188: 187 — 193.
- Krzymuski J. 1998. Postęp biologiczny w rolnictwie. W: Woś A. (red.). Encyklopedia agrobiznesu. Warszawa, Fundacja Innowacja — Wyższa Szkoła Społeczno-Ekonomiczna.
- Oleksiak T. 1998. Wpływ kwalifikowanego materiału siewnego na plonowanie zbóż. Biul. IHAR 208: 3 — 9.
- Oleksiak T. 2000. Pszenżyto w produkcji — wykorzystanie efektów hodowli. Zeszyty Naukowe Akademii Rolniczej w Szczecinie, 206 Agriculturae 82: 199 — 204.
- Oleksiak T. 2004 Rynek nasion. Rynek Środków Produkcji i Usług dla Rolnictwa, Nr 25: 36 — 42.
- Oleksiak T. 2006 Rynek nasion. Rynek Środków Produkcji i Usług dla Rolnictwa, Nr 30: 30 — 36.
- Rawlings J. O., Pantula S. G., Dickey D. A. 2001. Applied regression analysis — a research tool. 2nd Edition. New York, USA, Springer-Verlag Inc.
- SAS Institute Inc. 2004 a. BASE SAS 9.1.3 Procedures guide. Cary, NC, USA, SAS Publishing, SAS Institute Inc.
- SAS Institute Inc. 2004 b. SAS 9.1 Companion for Windows. Cary, NC, USA, SAS Publishing, SAS Institute Inc.
- SAS Institute Inc. 2004 c. SAS/STAT 9.1 User's Guide. Cary, NC, USA, SAS Publishing, SAS Institute Inc.
- Schwarz G. 1978. Estimating the dimension of a model. Annals of Statistics, 6: 461 — 464.
- Skarżyńska A., Sadowska J., Reinstein J. 1996. Koszty jednostkowe i dochodowość produkcji rolniczej w gospodarstwach indywidualnych w 1995 roku. Zagadnienia Ekonomiki Rolnej 4/5: 80 — 123.
- Stokes M. E., Davis C. S. 2000. Categorical Data Analysis Using the SAS System. 2nd Edition. New York, USA, SAS Publishing, SAS Institute Inc., John Wiley & Sons Inc.
- Witek T. (red.) 1981. Waloryzacja rolniczej przestrzeni produkcyjnej Polski według gmin. Puławy, IUNG.