

STANISŁAW MAŃKO

Uniwersytet Technologiczno-Przyrodniczy
Bydgoszcz

WYKORZYSTANIE METOD STATYSTYCZNO- -MATEMATYCZNYCH W ANALIZACH TABELARYCZNYCH NA PRZYKŁADZIE PRODUKCJI, KOSZTÓW I NADWYŻEK BEZPOŚREDNICH W CHOWIE BYDŁA MLECZNEGO W 2006 ROKU

Wstęp

Podstawą większości badań ekonomicznych opartych na danych empirycznych jest analiza wyników zestawionych postaci tzw. szeregów rozdzielczych w układzie tabelarycznym. Jeśli badania takie oparte są na danych masowych, wnioski sformułowane na ich podstawie nie powinny budzić zastrzeżeń. W wielu jednak przypadkach próby badawcze obejmują kilkadziesiąt lub kilkaset obiektów badawczych i wówczas może się okazać, że różnice między średnimi z poszczególnych klas są mniejsze od różnic pomiędzy obserwacjami w obrębie klas. W takiej sytuacji wnioski dotyczące zaobserwowanych różnic mogą być nieuprawnione. Aby stwierdzone różnice można uznać za istotne, konieczne jest wsparcie zestawień tabelarycznych metodami analizy statystyczno-matematycznej. W przypadku badań biometrycznych, np. z zakresu doświadczalnictwa rolniczego, wnioskowanie z wykorzystaniem statystyki matematycznej jest praktyką codzienną. Wydaje się, że w badaniach ekonomiczno-rolniczych również należy zwrócić większą uwagę na zastosowanie tego typu metod.

Od kilku lat w Instytucie Ekonomiki Rolnictwa i Gospodarki Żywnościowej – PIB prowadzone są badania w ramach Systemu Zbierania Danych o Produktach Rolniczych AGROKOSZTY [4]. Podstawowym celem tych badań jest ocena zmian w poziomie produkcji i nakładów na podstawie wiarygodnych danych źródłowych oraz zróżnicowania nadwyżki bezpośredniej wybranych działalności produkcyjnych w gospodarstwach konwencjonalnych i ekologicznych. Ze względu na dużą pracochłonność i kosztowność badań terenowych, wielkość prób badawczych obejmuje najczęściej po sto kilkadziesiąt przypadków (obserwacji) dla poszczególnych działalności roślinnych i zwierzęcych. Do 2005 roku wyniki opracowywane były głównie z wykorzystaniem zestawień tabelarycznych. W analizie wyników za rok 2006 po raz pierwszy zastosowano badania statystyczno-matematyczne na przykładzie analizy kosztów i nadwyżek bezpośrednich dwóch działalności roślinnych, tj. pszenicy ozimej

i żyta [4, s. 204-225]. Wyniki tych badań potwierdzają, że zróżnicowanie poszczególnych analizowanych wielkości jest tak duże, iż różnice międzygrupowe są istotne tylko w przypadku niektórych analizowanych zmiennych.

Jedną z podstawowych działalności produkcyjnych analizowanych w ramach badań AGROKOSZTY jest chów bydła mlecznego. W 2006 r. zgromadzono dane ze 159 gospodarstw utrzymujących krowy, a wyniki analizowano w zależności od poziomu nadwyżki bezpośredniej, przeliczonej na 1 krowę rocznie, oraz w ujęciu regionalnym [4, s. 93]. W badaniach tych, analizując wartość produkcji, nakłady, koszty i nadwyżkę bezpośrednią w produkcji mleka, nie weryfikowano różnic za pomocą testów statystycznych.

Podstawowym celem opracowania jest ocena za pomocą metod statystyki matematycznej, na ile przyjęte zasady grupowania są wystarczające do oceny rzeczywistego zróżnicowania produkcji, kosztów i nadwyżek bezpośrednich oraz sprawdzenie możliwości wykorzystania parametrycznych i nieparametrycznych testów istotności różnic w tym zakresie.

W opracowaniu wykorzystano zasady grupowania wyników badań zastosowane przez zespół pracowników IERiGŻ-PIB pod kierunkiem A. Skarżyńskiej [4, s. 22-23]. W analizach statystycznych uwzględniono grupowanie według kwartyli nadwyżki bezpośredniej w przeliczeniu na krowę oraz według regionów FADN. Spośród bogatego zestawu szczegółowych zmiennych, charakteryzujących w programie AGROKOSZTY chów bydła mlecznego, wybrano tylko niektóre zmienne opisujące warunki i organizację produkcji, wydajność mleczną krów, wartość produkcji, ceny produktów, bezpośrednie koszty produkcji oraz nadwyżkę bezpośrednią (por. tab. 1). Ze względu na metodyczny charakter opracowania pominięto szereg zmiennych charakteryzujących szczegółowe składniki kosztów bezpośrednich dostępne w bazie AGROKOSZTY.

Metoda badań

Grupowanie danych według określonych cech ma sens wtedy, gdy stworzone grupy charakteryzują się większą jednorodnością w stosunku do całej próby badawczej, a średnie arytmetyczne zmiennych analitycznych obliczone dla poszczególnych grup różnią się istotnie między sobą. Podstawą analizy istotności różnic są parametryczne i nieparametryczne testy oparte na analizie wariancji [2, 3, 5].

Punktem wyjścia do badań z wykorzystaniem testów istotności różnic była wstępna ocena zróżnicowania wybranych zmiennych z wykorzystaniem średniej arytmetycznej, wartości minimalnej i maksymalnej, odchylenia standardowego oraz wskaźnika zmienności. W analizie rozkładów zmiennych wykorzystano nieparametryczny test Shapiro-Wilka. Test ten uznawany jest za najmocniejszy spośród testów stosowanych do weryfikacji normalności rozkładu zmiennych [1]. Hipoteza zerowa w tym teście brzmi: zmienna charakteryzuje się rozkładem normalnym. Odrzucenie hipotezy zerowej oznacza przyjęcie hipotezy alternatywnej: rozkład zmiennej odbiega od normalnego [1, 2, 3, 5]. Dodatkowo oceniono asymetrię zmiennych za pomocą wskaźnika skośności

oraz koncentrację za pomocą kurtozy. Im wskaźnik skośności i kurtoza bliższe są zeru, tym w mniejszym stopniu rozkład zmiennej odbiega od normalnego.

Podstawową metodą statystycznej analizy istotności różnic między grupami obiektów badawczych w przypadku zmiennych ilościowych jest analiza wariancji. Jednym z najmocniejszych i najczęściej stosowanych testów jest test F-Snedecora [2, 3, 5]. Test ten zaliczany jest do grupy testów parametrycznych, dających poprawne wyniki w przypadku zmiennych charakteryzujących się rozkładem normalnym. Test ten obliczono dla wszystkich analizowanych zmiennych, jednak w analizie wykorzystano tylko wyniki dla tych zmiennych, które charakteryzowały się rozkładem normalnym.

W badaniach ekonomicznych często mamy do czynienia ze zmiennymi, które nie mają rozkładu normalnego. W tym przypadku zachodzi konieczność stosowania testów nieparametrycznych. W sytuacji gdy zmienne mają charakter ilościowy, a liczba porównywanych grup jest większa od 2, możliwe jest zastosowanie testu H Kruskala-Wallisa [2, 3]. Za pomocą tego testu oceniono istotność różnic w stosunku do tych zmiennych, których rozkład odbiegał od normalnego.

Wyniki badań

Zestaw zmiennych oraz ich formalno-statystyczną charakterystykę przedstawiono w tabeli 1. Szczególnie wysokim zróżnicowaniem charakteryzują się zmienne opisujące wielkość obszarową gospodarstw, w których ewidencjonowano dane empiryczne (RO02, RO03, RO04). Stosunkowo mocno zróżnicowane jest także pogłowie krów (RO08), oszacowane koszty zużycia pasz nietowarowych (RO35) oraz koszty wymiany stada (RO22). Najmniejszym zróżnicowaniem charakteryzuje się cena mleka (RO10).

Podstawowe parametry charakteryzujące rozkład badanych cech przedstawiono w tabeli 2. Posługując się testem Shapiro-Wilka, można stwierdzić, że tylko 5 zmiennych spośród badanych charakteryzuje się rozkładem normalnym (poziom istotności testu $SW > 0,05$):

- wydajność mleczna krów (RO09),
- cena cieląt odsadzonych od krów (RO11),
- wartość produkcji ogółem (RO14),
- wartość mleka (RO15),
- nadwyżka bezpośrednia (RO45).

W przypadku pozostałych zmiennych hipotezę o rozkładzie normalnym należy odrzucić z prawdopodobieństwem błędu niższym od 5% ($p < 0,05$). Oznacza to, że oparcie analizy wariancji na teście F-Snedecora, mimo dużej odporności tego testu na odstępstwa zmiennych od rozkładu normalnego, może prowadzić do wyciągnięcia niewłaściwych wniosków.

Zbiorcze zestawienie analizy wariancji metodą Fishera-Snedecora dla grup wyodrębnionych według poziomu nadwyżki bezpośredniej, w przeliczeniu na 1 krowę, przedstawiono w tabeli 3.

Tabela 1

Wstępna charakterystyka różnicowania analizowanych zmiennych

| Nazwa zmiennej | Symbol | Liczba obserwacji | Średnia arytmetyczna | Wartość minimalna | Wartość maksymalna | Odchylenie standardowe | Wskaźnik zmienności |
|--|--------|-------------------|----------------------|-------------------|--------------------|------------------------|---------------------|
| Powierzchnia gruntów ornych (ha) | RO02 | 159 | 25,10 | 2,46 | 333,95 | 30,68 | 122,23 |
| Powierzchnia trwałych użytków zielonych (ha) | RO03 | 159 | 8,60 | 0,00 | 81,02 | 11,95 | 139,07 |
| Powierzchnia użytków rolnych (ha) | RO04 | 159 | 33,73 | 5,58 | 367,29 | 36,22 | 107,37 |
| Wskaźnik bonitacji użytków rolnych (pkt) | RO07 | 159 | 0,80 | 0,06 | 1,72 | 0,34 | 42,68 |
| Średni w roku stan krów mlecznych (szt.) | RO08 | 159 | 19,96 | 2,46 | 101,85 | 16,42 | 82,27 |
| Wydajność mleczna krów (l/szt.) | RO09 | 159 | 4853,40 | 1628,51 | 8452,27 | 1463,83 | 30,16 |
| Cena sprzedaży mleka (zł/l) | RO10 | 157 | 0,92 | 0,48 | 1,19 | 0,14 | 14,71 |
| Cena cieląt odsadzonych od krów (zł/kg) | RO11 | 159 | 9,02 | 3,77 | 15,46 | 2,11 | 23,37 |
| Cena sprzedaży krów wybrakowanych (zł/kg) | RO12 | 116 | 2,83 | 0,48 | 6,69 | 0,76 | 27,02 |
| Wartość produkcji ogółem (zł/szt.) | RO14 | 159 | 5284,43 | 1719,41 | 10650,09 | 1740,41 | 32,93 |
| w tym: wartość mleka (zł/szt.) | RO15 | 159 | 4575,12 | 1238,10 | 9930,39 | 1738,84 | 38,01 |
| Koszty bezpośrednie ogółem (zł/szt.) | RO21 | 159 | 2064,52 | 243,25 | 4679,04 | 755,68 | 36,60 |
| z tego: wymiana stada (zł/szt.) | RO22 | 159 | 346,33 | 0,00 | 1000,00 | 214,94 | 62,06 |
| Pasze z zakupu i własne towarowe (zł/szt.) | RO23 | 159 | 1170,55 | 159,67 | 3176,34 | 550,08 | 46,99 |
| Pasze nietowarowe (zł/szt.) | RO35 | 159 | 274,98 | 0,00 | 1097,81 | 193,64 | 70,42 |
| pozostałe koszty (zł/szt.) | RO40 | 159 | 272,66 | 32,52 | 951,02 | 158,94 | 58,29 |
| Nadwyżka bezpośrednia (zł/szt.) | RO45 | 159 | 3219,91 | -1595,64 | 6418,27 | 1454,86 | 45,18 |
| Nakłady pracy ogółem (RBH/szt.) | RO46 | 159 | 183,66 | 18,87 | 527,37 | 93,16 | 50,72 |
| w tym: nakłady pracy własnej (RBH/szt.) | RO47 | 159 | 176,41 | 18,87 | 527,37 | 95,08 | 53,90 |

Źródło: Opracowanie własne na podstawie danych AGROKOSZTY.

Tabela 2

Charakterystyka rozkładu analizowanych zmiennych

| Nazwa zmiennej | Symbol | Liczba obser- wacji | Test Shapiro- -Wilka | Poziom istotności testu SW | Skośność | Kurtoza |
|--|--------|------------------------|----------------------------|----------------------------------|----------|---------|
| Powierzchnia gruntów ornych (ha) | RO02 | 159 | 0,484 | 0,000 | 6,928 | 65,563 |
| Powierzchnia trwałych użytków zielonych (ha) | RO03 | 159 | 0,628 | 0,000 | 3,072 | 11,297 |
| Powierzchnia użytków rolnych (ha) | RO04 | 159 | 0,546 | 0,000 | 5,670 | 46,351 |
| Wskaźnik bonitacji użytków rolnych (pkt) | RO07 | 159 | 0,975 | 0,006 | 0,337 | -0,655 |
| Średni w roku stan krów mlecznych (szt.) | RO08 | 159 | 0,839 | 0,000 | 1,806 | 4,508 |
| Wydajność mleczna krów (l/szt.) | RO09 | 159 | 0,994 | 0,723 | 0,034 | -0,418 |
| Cena sprzedaży mleka (zł/l) | RO10 | 157 | 0,972 | 0,003 | -0,511 | 0,597 |
| Cena cieląt odsadzonych od krów (zł/kg) | RO11 | 159 | 0,992 | 0,569 | 0,144 | 0,038 |
| Cena sprzedaży krów wybrakowanych (zł/kg) | RO12 | 116 | 0,913 | 0,000 | 0,515 | 5,882 |
| Wartość produkcji ogółem (zł/szt.) | RO14 | 159 | 0,991 | 0,437 | 0,204 | -0,148 |
| w tym: wartość mleka (zł/szt.) | RO15 | 159 | 0,986 | 0,120 | 0,250 | -0,273 |
| Koszty bezpośrednie ogółem (zł/szt.) | RO21 | 159 | 0,974 | 0,004 | 0,588 | 0,485 |
| z tego: wymiana stada (zł/szt.) | RO22 | 159 | 0,951 | 0,000 | 0,212 | 0,240 |
| pasze z zakupu i własne towarowe (zł/szt.) | RO23 | 159 | 0,945 | 0,000 | 0,999 | 1,407 |
| pasze nietowarowe (zł/szt.) | RO35 | 159 | 0,920 | 0,000 | 1,250 | 2,491 |
| pozostałe koszty (zł/szt.) | RO40 | 159 | 0,933 | 0,000 | 0,969 | 1,169 |
| Nadwyżka bezpośrednia (zł/szt.) | RO45 | 159 | 0,988 | 0,195 | -0,271 | -0,129 |
| Nakłady pracy ogółem (RBH/szt.) | RO46 | 159 | 0,950 | 0,000 | 0,791 | 0,451 |
| w tym: nakłady pracy własnej (RBH/szt.) | RO47 | 159 | 0,952 | 0,000 | 0,775 | 0,397 |

Źródło: Opracowanie własne na podstawie danych AGROKOSZTY.

Tabela 3
**Analiza istotności różnic (test F-Snedecora) dla danych pogrupowanych
według nadwyżek bezpośrednich**

| Nazwa zmiennej | Symbol | Stopnie swobodny dla czynnika | Średni kwadrat dla czynnika | Stopnie swobodny | Średni kwadrat dla błędu | Test F-Snedecora | Poziomostności |
|--|--------|-------------------------------|-----------------------------|------------------|--------------------------|------------------|----------------|
| Powierzchnia gruntów ornych (ha) | RO02 | 2 | 5247 | 156 | 886 | 5,923 | 0,003 |
| Powierzchnia trwałych użytków zielonych (ha) | RO03 | 2 | 232 | 156 | 142 | 1,638 | 0,198 |
| Powierzchnia użytków rolnych (ha) | RO04 | 2 | 7567 | 156 | 1232 | 6,144 | 0,003 |
| Wskaźnik bonitacji użytków rolnych (pkt) | RO07 | 2 | 0,246 | 156 | 0,116 | 2,126 | 0,123 |
| Średni w roku stan krów mlecznych (szt.) | RO08 | 2 | 5264 | 156 | 206 | 25,599 | 0,000 |
| Wydajność mleczna krów (l/szt.) | RO09 | 2 | 105951429 | 156 | 811912 | 130,496 | 0,000 |
| Cena sprzedaży mleka (zł/l) | RO10 | 2 | 0,570 | 154 | 0,011 | 50,461 | 0,000 |
| Cena cieląt odsadzonych od krów (zł/kg) | RO11 | 2 | 3,391 | 156 | 4,456 | 0,761 | 0,469 |
| Cena sprzedaży krów wybrakowanych (zł/kg) | RO12 | 2 | 0,001 | 113 | 0,593 | 0,003 | 0,997 |
| Wartość produkcji ogółem (zł/szt.) | RO14 | 2 | 167219455 | 156 | 924023 | 180,969 | 0,000 |
| w tym: wartość mleka (zł/szt.) | RO15 | 2 | 168038081 | 156 | 907982 | 185,068 | 0,000 |
| Koszty bezpośrednie ogółem (zł/szt.) | RO21 | 2 | 2190040 | 156 | 550298 | 3,980 | 0,021 |
| z tego: wymiana stada (zł/szt.) | RO22 | 2 | 182801 | 156 | 44448 | 4,113 | 0,018 |
| pasze z zakupu i własne towarowe (zł/szt.) | RO23 | 2 | 471937 | 156 | 224049 | 2,106 | 0,125 |
| pasze nietowarowe (zł/szt.) | RO35 | 2 | 139878 | 156 | 36185 | 3,866 | 0,023 |
| pozostałe koszty (zł/szt.) | RO40 | 2 | 255848 | 156 | 22304 | 11,471 | 0,000 |
| Nadwyżka bezpośrednia (zł/szt.) | RO45 | 2 | 133972620 | 156 | 426161 | 314,371 | 0,000 |
| Nakłady pracy ogółem (RBH/szt.) | RO46 | 2 | 73138 | 156 | 7852 | 9,315 | 0,000 |
| w tym: nakłady pracy własnej (RBH/szt.) | RO47 | 2 | 92343 | 156 | 7972 | 11,583 | 0,000 |

Źródło: Opracowanie własne na podstawie danych AGROKOSZTY.

W przypadku wszystkich zmiennych charakteryzujących się rozkładem normalnym należy odrzucić hipotezę o braku różnic pomiędzy średnimi dla wyłonionych grup danych z prawdopodobieństwem mniejszym od 0,001. Wprawdzie w przypadku niektórych innych zmiennych poziom istotności testu F-Snedecora jest niższy od 0,05, jednak wnioski co do istotności różnic należy formułować ostrożnie. Zgodnie z założeniami metodycznymi, w weryfikacji istotności różnic dla pozostałych zmiennych wykorzystano nieparametryczny test H Kruskala-Wallisa (tab. 4).

Z zestawienia testu H Kruskala-Wallisa (tab. 4) wynika, że klasyfikacja oparta na kwartylach nadwyżki bezpośredniej dała pozytywne rezultaty w stosunku do większości analizowanych zmiennych. Nieistotne okazały się różnice dotyczące:

- powierzchni użytków zielonych w gospodarstwie (RO03),
- wskaźnika bonitacji użytków rolnych (RO07),
- ceny sprzedaży krów wybrakowanych (RO12).

Na pograniczu istotności różnic, lecz powyżej granicznej wielkości ($p=0,05$), znalazły się zmienne:

- koszty bezpośrednie ogółem (RO21),
- pasze z zakupu i własne towarowe (RO23),
- pasze nietowarowe (RO35).

Można więc stwierdzić, że w przypadku bydła mlecznego stosowana w ramach programu AGROKOSZTY klasyfikacja danych oparta na kwartylach wartości nadwyżki bezpośredniej dała poprawne wyniki. Przy praktycznym wykorzystywaniu danych dotyczących produkcji, kosztów i nadwyżek bezpośrednich należy także zwrócić uwagę na wielkość obszarową gospodarstw, wielkość pogłowia krów oraz wydajność mleczną. Wydaje się, że cechy te powinny być dodatkowo uwzględniane w klasyfikacji danych charakteryzujących chów bydła mlecznego.

Drugim sposobem prezentowania wyników badań w ramach programu AGROKOSZTY jest grupowanie według położenia gospodarstw w regionach FADN [4]. Wyniki analizy istotności różnic dla tak sklasyfikowanych danych przedstawiono w tabeli 5.

Analiza zmiennych, dla których stwierdzono rozkład normalny (RO09, RO11, RO14, RO15, RO45), pozwala stwierdzić, że tylko w przypadku ceny cieląt (RO11) można odrzucić hipotezę o braku różnic między średnimi w podgrupach (tab. 5). W przypadku pozostałych 4 zmiennych brak jest podstaw, że występują różnice regionalne.

Analiza istotności różnic metodą H Kruskala-Wallisa wskazuje na istotne różnice dotyczące wielkości obszarowej gospodarstw (RO02, RO03, RO04), jakości gleb (RO07), wielkości stada krów w gospodarstwach (RO08), wysokości kosztów bezpośrednich w przeliczeniu na krowę (RO21), a także ich elementów składowych: kosztów pasz z zakupu i własnych potencjalnie towarowych (RO23), pasz własnych nietowarowych (RO35) i pozostałych kosztów bezpośrednich (RO40) oraz nakładów pracy (RO46, RO47). Brak jest natomiast podstaw do stwierdzenia, że istnieją regionalne różnice w zakresie wydajności mlecznej krów (RO09), cen mleka (RO10) i wybrakowanych krów

(RO12) oraz wartości nadwyżki bezpośredniej (RO45). Mimo więc znacznych różnic w zakresie warunków i organizacji produkcji mleka oraz stosowanych technologii produkcji w poszczególnych regionach, różnice w poziomie wyników ekonomicznych mierzonych nadwyżką bezpośrednią w przeliczeniu na 1 krowę są znikome. Różnice te nie wpływają także na poziom mleczności krów.

Tabela 4

Analiza istotności różnic (test H Kruskala-Wallisa) dla danych pogrupowanych według nadwyżek bezpośrednich

| Nazwa zmiennej | Symbol | Stopnie swobody | Liczba obserwacji | Test H Kruskala-Wallisa | Poziom istotności |
|--|--------|-----------------|-------------------|-------------------------|-------------------|
| Powierzchnia gruntów ornyc (ha) | RO02 | 2 | 159 | 20,843 | 0,000 |
| Powierzchnia trwałych użytków zielonych (ha) | RO03 | 2 | 159 | 3,859 | 0,145 |
| Powierzchnia użytków rolnych (ha) | RO04 | 2 | 159 | 19,640 | 0,000 |
| Wskaźnik bonitacji użytków rolnych (pkt) | RO07 | 2 | 159 | 3,051 | 0,218 |
| Średni w roku stan krów mlecznych (szt.) | RO08 | 2 | 159 | 56,661 | 0,000 |
| Wydajność mleczna krów (l/szt.) | RO09 | 2 | 159 | 99,860 | 0,000 |
| Cena sprzedaży mleka (zł/l) | RO10 | 2 | 157 | 60,547 | 0,000 |
| Cena cieląt odsadzonych od krów (zł/kg) | RO11 | 2 | 159 | 1,177 | 0,555 |
| Cena sprzedaży krów wybrakowanych (zł/kg) | RO12 | 2 | 116 | 0,228 | 0,892 |
| Wartość produkcji ogółem (zł/szt.) | RO14 | 2 | 159 | 113,377 | 0,000 |
| w tym: wartość mleka (zł/szt.) | RO15 | 2 | 159 | 114,048 | 0,000 |
| Koszty bezpośrednie ogółem (zł/szt.) | RO21 | 2 | 159 | 5,859 | 0,053 |
| z tego: wymiana stada (zł/szt.) | RO22 | 2 | 159 | 10,492 | 0,005 |
| pasze z zakupu i własne towarowe (zł/szt.) | RO23 | 2 | 159 | 5,620 | 0,060 |
| pasze nietowarowe (zł/szt.) | RO35 | 2 | 159 | 5,835 | 0,054 |
| pozostałe koszty (zł/szt.) | RO40 | 2 | 159 | 18,185 | 0,000 |
| Nadwyżka bezpośrednia (zł/szt.) | RO45 | 2 | 159 | 133,594 | 0,000 |
| Nakłady pracy ogółem (RBH/szt.) | RO46 | 2 | 159 | 17,297 | 0,000 |
| w tym: nakłady pracy własnej (RBH/szt.) | RO47 | 2 | 159 | 21,797 | 0,000 |

Źródło: Opracowanie własne na podstawie danych AGROKOSZTY.

Tabela 5

**Analiza istotności różnic (test F-Snedecora) dla danych pogrupowanych
według położenia regionalnego**

| Nazwa zmiennej | Symbol | Stopnie swobodny dla czynnika | Średni kwadrat dla czynnika | Stopnie swobodny dla błęd | Średni kwadrat dla błęd | Test F-Snedecora | Poziomostwo |
|--|--------|-------------------------------|-----------------------------|---------------------------|-------------------------|------------------|-------------|
| Powierzchnia gruntów ornych (ha) | RO02 | 3 | 3812 | 155 | 886 | 4,304 | 0,006 |
| Powierzchnia trwałych użytków zielonych (ha) | RO03 | 3 | 727 | 155 | 132 | 5,524 | 0,001 |
| Powierzchnia użytków rolnych (ha) | RO04 | 3 | 7018 | 155 | 1201 | 5,842 | 0,001 |
| Wskaźnik bonitacji użytków rolnych (pkt) | RO07 | 3 | 0,632 | 155 | 0,107 | 5,890 | 0,001 |
| Średni w roku stan krów mlecznych (szt.) | RO08 | 3 | 1000 | 155 | 256 | 3,913 | 0,010 |
| Wydajność mleczna krów (l/szt.) | RO09 | 3 | 1869238 | 155 | 2148086 | 0,870 | 0,458 |
| Cena sprzedaży mleka (zł/l) | RO10 | 3 | 0,008 | 153 | 0,019 | 0,416 | 0,742 |
| Cena cieląt odsadzonych od krów (zł/kg) | RO11 | 3 | 28,516 | 155 | 3,977 | 7,171 | 0,000 |
| Cena sprzedaży krów wybrakowanych (zł/kg) | RO12 | 3 | 0,421 | 112 | 0,588 | 0,716 | 0,544 |
| Wartość produkcji ogółem (zł/szt.) | RO14 | 3 | 1784282 | 155 | 3053120 | 0,584 | 0,626 |
| w tym: wartość mleka (zł/szt.) | RO15 | 3 | 2997255 | 155 | 3024062 | 0,991 | 0,399 |
| Koszty bezpośrednie ogółem (zł/szt.) | RO21 | 3 | 2593154 | 155 | 531917 | 4,875 | 0,003 |
| z tego: wymiana stada (zł/szt.) | RO22 | 3 | 11955 | 155 | 46862 | 0,255 | 0,858 |
| pasze z zakupu i własne towarowe (zł/szt.) | RO23 | 3 | 1175248 | 155 | 285695 | 4,114 | 0,008 |
| pasze nietowarowe (zł/szt.) | RO35 | 3 | 203289 | 155 | 34289 | 5,929 | 0,001 |
| pozostałe koszty (zł/szt.) | RO40 | 3 | 65324 | 155 | 24485 | 2,668 | 0,050 |
| Nadwyżka bezpośrednia (zł/szt.) | RO45 | 3 | 364305 | 155 | 2150538 | 0,169 | 0,917 |
| Nakłady pracy ogółem (RBH/szt.) | RO46 | 3 | 44194 | 155 | 7991 | 5,531 | 0,001 |
| w tym: nakłady pracy własnej (RBH/szt.) | RO47 | 3 | 55632 | 155 | 8138 | 6,836 | 0,000 |

Źródło: Opracowanie własne na podstawie danych AGROKOSZTY.

Tabela 6

**Analiza istotności różnic (test H Kruskala-Wallisa) dla danych pogrupowanych
według położenia regionalnego**

| Nazwa zmiennej | Symbol | Stopnie swobody | Liczba obserwacji | Test H Kruskala-Wallisa | Poziom istotności |
|--|--------|-----------------|-------------------|-------------------------|-------------------|
| Powierzchnia gruntów orných (ha) | RO02 | 3 | 159 | 25,362 | 0,000 |
| Powierzchnia trwałych użytków zielonych (ha) | RO03 | 3 | 159 | 16,754 | 0,001 |
| Powierzchnia użytków rolnych (ha) | RO04 | 3 | 159 | 20,187 | 0,000 |
| Wskaźnik bonitacji użytków rolnych (pkt) | RO07 | 3 | 159 | 15,152 | 0,002 |
| Średni w roku stan krów mlecznych (szt.) | RO08 | 3 | 159 | 10,950 | 0,012 |
| Wydajność mleczna krów (l/szt.) | RO09 | 3 | 159 | 2,940 | 0,401 |
| Cena sprzedaży mleka (zł/l) | RO10 | 3 | 157 | 2,451 | 0,484 |
| Cena cieląt odsadzonych od krów (zł/kg) | RO11 | 3 | 159 | 16,942 | 0,001 |
| Cena sprzedaży krów wybrakowanych (zł/kg) | RO12 | 3 | 116 | 3,371 | 0,338 |
| Wartość produkcji ogółem (zł/szt.) | RO14 | 3 | 159 | 1,614 | 0,656 |
| w tym: wartość mleka (zł/szt.) | RO15 | 3 | 159 | 2,921 | 0,404 |
| Koszty bezpośrednie ogółem (zł/szt.) | RO21 | 3 | 159 | 16,356 | 0,001 |
| z tego: wymiana stada (zł/szt.) | RO22 | 3 | 159 | 0,717 | 0,869 |
| pasze z zakupu i własne towarowe (zł/szt.) | RO23 | 3 | 159 | 28,881 | 0,000 |
| pasze nietowarowe (zł/szt.) | RO35 | 3 | 159 | 16,689 | 0,001 |
| pozostałe koszty (zł/szt.) | RO40 | 3 | 159 | 9,823 | 0,020 |
| Nadwyżka bezpośrednia (zł/szt.) | RO45 | 3 | 159 | 0,305 | 0,959 |
| Nakłady pracy ogółem (RBH/szt.) | RO46 | 3 | 159 | 16,422 | 0,001 |
| w tym: nakłady pracy własnej (RBH/szt.) | RO47 | 3 | 159 | 21,443 | 0,000 |

Źródło: Opracowanie własne na podstawie danych AGROKOSZTY.

Podsumowanie

Najczęściej stosowaną metodą opracowywania danych empirycznych w badaniach ekonomiczno-rolniczych są zestawienia tabelaryczne. Jest to sposób bardzo prosty i w wypadku analizy masowych danych pozwala na formułowanie poprawnych wniosków. W wielu jednak wypadkach, ze względu na duże zróżnicowanie danych w obrębie wyodrębnionych klas, wnioski mogą być obciążone znacznym błędem. Zespół badawczy pod kierunkiem A. Skarżyńskiej, analizując wyniki badań za 2006 r. w ramach programu AGROKOSZTY, postanowił wykorzystać bardziej precyzyjne metody analizy danych empirycznych. Na przykładzie dwóch roślinnych działalności produkcyjnych, oprócz tradycyjnych zasad analizy, zastosowano metody statystyki matematycznej. Wyniki analizy wskazywały na celowość stosowania metod z tego zakresu [4]. W ramach tego projektu dotychczas nie analizowano jednak działalności zwierzęcych. W opracowaniu, korzystając z bazy danych AGROKOSZTY, dotyczących produkcji, kosztów i nadwyżek bezpośrednich w chowie bydła mlecznego, podjęto próbę zweryfikowania przydatności metod statystyczno-matematycznych do oceny poprawności wnioskowania na podstawie zestawień tabelarycznych.

Z analizy formalno-statystycznej wynika, że większość zmiennych opisujących wyniki produkcyjne, koszty, nadwyżki bezpośrednie oraz warunki i organizację produkcji mleka charakteryzuje się rozkładem odbiegającym od normalnego. Stosowanie metod analizy wariancji opartych na testach parametrycznych w tym przypadku ma ograniczoną przydatność. Podstawową metodą weryfikacji istotności różnic powinno być stosowanie testów nieparametrycznych. W przypadku zmiennych ilościowych i liczbie klas większej od dwóch przydatny może być test H Kruskala-Wallisa.

Zastosowanie testu F-Snedecora w odniesieniu do zmiennych o rozkładzie normalnym i H Kruskala-Wallisa w stosunku do pozostałych zmiennych pozwoliło stwierdzić, że większość wniosków sformułowanych przez A. Skarżyńską [4, s. 98-103], dotyczących zróżnicowania poszczególnych wielkości charakteryzujących chów bydła mlecznego na tle klas wyłonionych na podstawie poziomu nadwyżek bezpośrednich, jest poprawna. Pewne wątpliwości mogą tylko budzić wnioski dotyczące kosztów bezpośrednich, których zróżnicowanie w świetle zastosowanych testów statystycznych okazało się nieistotne, aczkolwiek na pograniczu dopuszczalnego błędu statystycznego. Uwzględniając istotność różnic w obrębie zmiennych charakteryzujących warunki, skalę produkcji oraz wydajność mleczną krów na tle zróżnicowania nadwyżek bezpośrednich, wydaje się, że celowe byłoby uwzględnienie także wpływu skali produkcji i wydajności mlecznej krów na wielkość nadwyżek bezpośrednich. W tym celu warto byłoby zastosować dodatkowe kryteria klasyfikacji danych lub posłużyć się opisowymi modelami ekonometrycznymi.

Wnioski dotyczące zróżnicowania regionalnego [4, s. 103-113] w zasadzie także nie budzą większych zastrzeżeń, jednak w świetle analizy statystyczno-matematycznej regionalne różnice w wydajności mlecznej krów nie mogą być uznane za istotne.

W celu zwiększenia poziomu ufności wnioskowania w ramach badań AGROKOSZTY wydaje się uzasadnione w szerszym zakresie wsparcie analizy metodami z zakresu statystyki matematycznej. Jest to szczególnie ważne ze względu na niezbyt liczne próby badawcze gromadzonych danych dla poszczególnych działalności produkcyjnych. Wybór szczegółowych testów powinien uwzględniać charakter rozkładu analizowanych zmiennych. Ze względu na fakt, że większość zmiennych nie spełnia warunku rozkładu normalnego, należy przede wszystkim posługiwać się testami nieparametrycznymi.

Literatura:

1. Domański C.: Statystyczne testy nieparametryczne. PWE, Warszawa 1979.
2. Mynarski S.: Analiza danych rynkowych i marketingowych z wykorzystaniem programu STATISTICA. Wydawnictwo AE w Krakowie, Kraków 2003.
3. Mynarski S.: Praktyczne metody analizy danych rynkowych i marketingowych. Kantor Wydawniczy Zakamycze 2000.
4. Produkcja, koszty i nadwyżka bezpośrednia wybranych produktów rolniczych w 2006 roku. Praca zbiorowa pod red. A. Skarżyńskiej. PW nr 60, IERiGŻ, Warszawa 2007.
5. Steczkowski J., Zeliaś A.: Analiza wariancyjna i kowariancyjna w badaniach ekonomicznych. PWE, Warszawa 1982.

STANISŁAW MAŃKO

University of Technology and Life Sciences
Bydgoszcz

USE OF STATISTICAL AND MATHEMATICAL METHODS IN TABLE ANALYSIS ON THE EXAMPLE OF THE PRODUCTION, COSTS AND DIRECT SURPLUSES IN DAIRY CATTLE PRODUCTION IN 2006

Summary

The Paper addresses the problem of the purposefulness of using statistical and mathematical methods of the analysis of production, costs and direct surpluses. The research shows that the application of variance analysis may increase the level of reliability of conclusions, it does not, however, question most conclusions based on the analysis of table lists. The selection of detailed tests should take account of the nature of the distribution of analysed variables. Since most variables do not meet the conditions of normal distribution, non-parameter tests should be used in most cases.