

ZASTOSOWANIE POSTĘPOWANIA SEKWENCYJNEGO W ANALIZIE CECH GEOMETRYCZNYCH NASION WYBRANYCH ROŚLIN ZBOŻOWYCH

Dorota Domagała

Katedra Zastosowań Matematyki i Informatyki, Uniwersytet Przyrodniczy w Lublinie

Tomasz Guz

Katedra Inżynierii i Maszyn Spożywczych, Uniwersytet Przyrodniczy w Lublinie

Streszczenie. Badaniami objęte zostały nasiona wybranych roślin zbożowych (jęczmień, owies, pszenica, pszenżyto, żyto) o wilgotności 14, 16 i 18%. Dokonano pomiaru podstawowych cech geometrycznych za pomocą systemu komputerowej analizy obrazu. Do analizy wpływu wilgotności na rzutowe pole powierzchni i długość ziaren zastosowano testy sekwencyjne.

Słowa kluczowe: nasiona zbóż, komputerowa analiza obrazów, test sekwencyjny

Wstęp

W klasycznej teorii weryfikacji hipotez statystycznych liczebność próby jest stała i określana w momencie planowania eksperymentu. Dopóki nie zostanie wylosowany ostatni element informacji uzyskane z już wylosowanych elementów nie powodują podjęcia żadnej decyzji odnośnie weryfikowanej hipotezy. Może zdarzyć się, że wylosujemy zbyt dużo lub zbyt mało elementów do próby. Wnioskowanie statystyczne na podstawie małych prób jest często obarczone dużym ryzykiem podjęcia błędnej decyzji odnośnie postawionej hipotezy statystycznej. Z kolei zbyt liczna próba może znacznie podnieść koszt eksperymentu. W wielu badaniach, np. w medycynie, biologii, kontroli jakości i wielu innych naukach, uzyskanie wystarczająco dużych prób jest bardzo trudne lub nawet niemożliwe. W takich sytuacjach alternatywą staje się analiza sekwencyjna.

W odróżnieniu od podejścia klasycznego, w podejściu sekwencyjnym liczebność próby jest zmienną losową. Fisz [1958] określa postępowanie sekwencyjne przy weryfikacji hipotezy statystycznej jako takie, w którym liczebność próby nie jest dana z góry, a wartość badanej cechy każdego wylosowanego elementu może spowodować przyjęcie weryfikowanej hipotezy lub jej odrzucenie, lub też kontynuowanie badania. Dane potrzebne w celu zweryfikowania hipotezy statystycznej są uzyskiwane w trakcie przeprowadzania badania, a badacz już od początku eksperymentu stosuje statystykę. Podejście sekwencyjne umożliwia zatem stałą kontrolę nad przebiegiem eksperymentu [Marek, Noworol 1987].

Ogólnie procedura sekwencyjna dotycząca weryfikowania hipotez składa się z następujących kroków:

1. Określamy prawdopodobieństwo α popełnienia błędu, polegającego na odrzuceniu hipotezy w przypadku, gdy jest prawdziwa (błąd I rodzaju) oraz prawdopodobieństwo β popełnienia błędu, polegającego na przyjęciu hipotezy w przypadku, gdy jest fałszywa (błąd II rodzaju).
2. Obliczamy wartość odpowiedniej statystyki testowej I_k , nazywanej także ilorazem wiarygodności. Małe wartości statystyki I_k świadczą na korzyść hipotezy zerowej, zaś duże na korzyść hipotezy alternatywnej. Jeśli wartość statystyki I_k jest mniejsza od pewnej stałej A , to badanie kończy się przyjęciem hipotezy zerowej, jeśli większa od pewnej stałej B , to badanie kończy się przyjęciem hipotezy alternatywnej. W sytuacji, gdy $A < I_k < B$ wówczas kontynuujemy pobieranie próby. Stałe A i B są zależne od ustalonych wartości α i β w następujący sposób:

$$A = \frac{\beta}{1-\alpha}, \quad B = \frac{1-\beta}{\alpha} \quad \text{ i } 0 < A < 1 < B.$$

Jak widać, w odróżnieniu od podejścia klasycznego w podejściu sekwencyjnym nastawiamy się nie na odrzucenie hipotezy, lecz na jej przyjęcie. W przypadku prawdziwości hipotezy zerowej postępowanie sekwencyjne pozwala na podjęcie decyzji o przyjęciu tej hipotezy, natomiast postępowanie klasyczne pozwala wnioskować jedynie o braku podstaw do jej odrzucenia. W przypadku prawdziwości hipotezy alternatywnej ujęcie sekwencyjne prowadzi do jej przyjęcia, zaś ujęcie klasyczne do odrzucenia hipotezy zerowej [Marek, Noworol 1987].

Kształt oraz wymiary ziarna zależą od własności odmianowych, klimatu, warunków agrotechnicznych oraz warunków wzrostu roślin [Łukaszuk 2005]. Istnieje ścisła zależność między kształtem ziarna, a jego wymiarami. Kształt ziarniaków może być kulisty, jajowaty, owalny lub klinowaty. Może ponadto występować wyraźna strona grzbietowa i brzuszna oraz podstawa i wierzchołek. Każde ziarno charakteryzuje się parametrami takimi jak długość, szerokość oraz grubość [Ciećko 2003]. Zastosowanie nowych urządzeń pomiarowych (systemy komputerowej analizy obrazu) pozwalają obliczyć jeszcze bardziej złożone parametry obiektów biologicznych oparte na pomiarze pola powierzchni, długości linii brzegowych czy też średnic zastępczych [Zdunek 2007].

Cel i zakres pracy

Celem pracy było zweryfikowanie hipotezy o braku wpływu poziomu wilgotności ziaren wybranych gatunków roślin zbożowych na wybrane cechy geometryczne: rzutowe pole powierzchni i długość ziarna za pomocą analizy sekwencyjnej.

Materiał badań

Ziarna pięciu zbóż: żyta odmiany Słowiańskie, owsa odmiany Sławko, pszenicy odmiany Tonacja, pszenżyta odmiany Pawo oraz jęczmienia odmiany Stratus zostały dowilżone

do wilgotności 14, 16 oraz 18%. Z tak przygotowanego surowca pobierano losowo próbki ziaren. Ziarna te układano bruzdką na powierzchni stolika przedmiotowego podświetlanego od spodu białym światłem o stałej temperaturze barwowej 5500K. Fotografie ziaren zapisywano w komputerze za pomocą programu do analizy obrazów SUPEVIST. Obszar obserwacji był skalowany w celu uzyskania wartości długości i pola powierzchni rzutowych ułożonych w ten sposób ziaren. Fotografie cyfrowe uzyskanych obiektów poddano filtracji w celu wyodrębnienia ich od tła. W tym celu wykonano progowanie obrazów pierwotnych i przekształcenie ich w obrazy binarne, które posłużyły jako matryce pomiarowe. Wykorzystując możliwości pomiarowe programu SVISTMET uzyskano rzutowe pole powierzchni i długość ziarna.

Opis metody

Zakładamy, że mamy l populacji, w których badana cecha ma rozkład normalny ze średnimi $\mu_1, \mu_2, \dots, \mu_l$, odpowiednio, i wspólną wariancją σ^2 . Weryfikujemy hipotezę zerową o równości wszystkich średnich

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_l$$

przeciwko hipotezie alternatywnej

$$H_1 : \text{nie wszystkie wartości średnie są równe.}$$

W sekwencyjnej procedurze dla jednoczynnikowej analizy wariancji na każdym etapie pobierana jest ustalona liczba n_k obserwacji z każdej populacji. Niech x_{ijk} oznacza j -tą obserwację w i -tej populacji na k -tym etapie ($i = 1, \dots, l, j = 1, \dots, n_k, k = 1, 2, \dots$). Na każdym etapie procedury sekwencyjnej, po każdym pobraniu elementu do próby prowadzimy analizę wariancji o ustalonej wielkości próby $N_k = ln_k$. Wyliczamy kolejno, jak w klasycznej analizie wariancji [Krysicki i in. 1998]:

sumę kwadratów odchyłeń między grupami

$$SS_{mg} = n_k \sum_{i=1}^l (\bar{x}_{i\bullet k} - \bar{x}_{\bullet\bullet k})^2,$$

gdzie

$$\bar{x}_{i\bullet k} = \frac{1}{n_k} \sum_{j=1}^{n_k} x_{ijk}, \quad \bar{x}_{\bullet\bullet k} = \frac{1}{n_k l} \sum_{i=1}^l \sum_{j=1}^{n_k} x_{ijk}$$

sumę kwadratów odchyłeń wewnątrz grup

$$SS_{wg} = \sum_{i=1}^l \sum_{j=1}^{n_k} (x_{ijk} - \bar{x}_{i\bullet k})^2$$

średni kwadrat odchyłeń między grupami

$$MS_{mg} = \frac{SS_{mg}}{l-1}$$

średni kwadrat odchyłeń wewnątrz grup

$$MS_{wg} = \frac{SS_{wg}}{l(n_k - 1)}$$

oraz

$$F_k = \frac{MS_{mg}}{MS_{wg}}$$

i wartość statystyki testowej I_k określonej poniższym wzorem [Marek, Noworol 1987]

$$I_k = \frac{1}{(1+k\lambda)^{(l-1)/2}} \left[\frac{1 + \frac{l-1}{lk-1} F_k}{1 + \frac{l-1}{lk-1} F_k} \right]^{(lk-1)/2}$$

gdzie

$$\lambda = \sqrt{\frac{n_k}{(l-1)\sigma^2} \sum_{i=1}^l (\mu_i - \mu)^2}, \quad \mu = \frac{1}{l} \sum_{i=1}^l \mu_i.$$

Statystykę I_k można przedstawiać także w postaci logarytmicznej [Siegmond 1980].

Wyniki i dyskusja

Analizę przeprowadzono osobno zarówno dla poszczególnych gatunków zbóż jak i dla obydwu cech. Zatem porównywano rzutowe pole powierzchni oraz długość ziarna w trzech populacjach ziaren danego gatunku o wilgotności odpowiednio 14, 16 i 18%. Analizę statystyczną przeprowadzono dla ustalonej wartości $\beta = 0,1$ i $\beta = 0,2$ oraz $\alpha = 0,01$ i $\alpha = 0,05$. W etapie pierwszym postępowania sekwencyjnego z każdej z trzech populacji pobrano próbki o liczebności $n_1 = 2$ i obliczano wartość I_k . Jeżeli na danym etapie postępowania nie było możliwe przyjęcie hipotezy zerowej bądź alternatywnej (spełniona była nierówność $A < I_k < B$), to w kolejnym etapie zwiększano liczebność próbek o 1, ale tylko do momentu osiągnięcia przez n_k wartości 100. Uzyskane wyniki – informację o przyjęciu danej hipotezy oraz liczebność próby, na podstawie której podjęto taką decyzję przedstawia tabela 1 dla rzutowego pola powierzchni ziarna oraz tabela 2 dla długości ziarna.

Zastosowanie postępowania sekwencyjnego...

Tabela 1. Rzutowe pole powierzchni ziarna - decyzja o przyjęciu hipotezy H_0 lub H_1 przy liczebności próbki n_k
 Table 1. Projective surface area of grain - decision on acceptance of hypothesis H_0 or H_1 for sample size n_k

Wyszczególnienie		$\beta = 0,1$		$\beta = 0,2$	
		$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,01$
Jęczmień	decyzja (przyjęcie hipotezy)	H_0	H_0	H_0	H_0
	liczebność n_k	29	29	12	12
Owies	decyzja (przyjęcie hipotezy)	H_1	H_1	H_1	H_1
	liczebność n_k	13	14	13	14
Pszenica	decyzja (przyjęcie hipotezy)	H_0	H_0	H_0	H_0
	liczebność n_k	24	24	12	14
Pszennyto	decyzja (przyjęcie hipotezy)	H_1	H_1	H_0	H_0
	liczebność n_k	57	69	13	13
Żyto	decyzja (przyjęcie hipotezy)	H_1	H_1	H_1	H_1
	liczebność n_k	9	10	9	10

Źródło: obliczenia własne autorów

Dla ziaren jęczmienia i pszenicy przyjęto hipotezę zerową o równości wszystkich średnich, czyli poziom wilgotności ziarna nie miał wpływu na wartość rzutowego pola powierzchni zarówno dla $\beta = 0,1$ i $\beta = 0,2$. Dla owsa i żyta stwierdzono istotny wpływ wilgotności ziarna na wartość rzutowego pola powierzchni w przypadku obydwu rozpatrywanych wartości β . Natomiast w przypadku pszenżyta dla $\beta = 0,2$ podjęto decyzję o przyjęciu hipotezy zerowej a dla $\beta = 0,1$ - decyzję o przyjęciu hipotezy alternatywnej.

Tabela 2. Długość ziarna - decyzja o przyjęciu hipotezy H_0 lub H_1 przy liczebności próbki n_k
 Table 2. Grain length - decision on acceptance of hypothesis H_0 or H_1 for sample size n_k

Wyszczególnienie		$\beta = 0,1$		$\beta = 0,2$	
		$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,01$
Jęczmień	decyzja (przyjęcie hipotezy)	H_0	H_0	H_0	H_0
	liczebność n_k	45	46	20	20
Owies	decyzja (przyjęcie hipotezy)	H_1	H_1	H_1	H_1
	liczebność n_k	12	13	12	13
Pszenica	decyzja (przyjęcie hipotezy)	H_0	brak	H_0	H_0
	liczebność n_k	24	100	12	13
Pszennyto	decyzja (przyjęcie hipotezy)	H_1	brak	H_0	H_0
	liczebność n_k	72	100	11	11
Żyto	decyzja (przyjęcie hipotezy)	H_1	H_1	H_1	H_1
	liczebność n_k	46	47	46	47

Źródło: obliczenia własne autorów

W przypadku długości ziarna decyzję o przyjęciu hipotezy zerowej podjęto dla jęczmienia (wszystkie rozpatrywane wartości α oraz β) i pszenicy (poza sytuacją, gdy $\alpha = 0,01$ i $\beta = 0,1$ - wtedy próba o liczebności $n_{99} = 100$ nie doprowadziła do przyjęcia żadnej z hipotez). Dla ziaren owsa i żyta otrzymujemy wnioski jak dla rzutowego pola powierzchni. W przypadku pszenżyta przyjęto hipotezę zerową dla $\beta = 0,2$, natomiast hipotezę alternatywną dla $\beta = 0,1$ i $\alpha = 0,01$. W sytuacji, gdy $\beta = 0,1$ i $\alpha = 0,01$ nie podjęto decyzji o przyjęciu żadnej z hipotez – należałoby kontynuować pobieranie próby.

Wnioski

1. Dla rzutowego pola powierzchni decyzję przyjęcia odpowiedniej hipotezy, zerowej bądź alternatywnej, podjęto w przypadku każdego gatunku zboża.
2. Dla długości ziarna, w przypadku pszenicy i pszenżyta, ostatni przeprowadzony etap postępowania sekwencyjnego oparty na 100 elementowych próbkach przyniósł informację o konieczności kontynuowania pobierania próby. Ta sytuacja wystąpiła dla $\alpha = 0,01$ i $\beta = 0,1$. Dla pozostałych gatunków zbóż przyjęto hipotezę o równości średnich bądź o istnieniu między nimi istotnych statystycznie różnic.
3. Decyzję przyjęcia odpowiedniej hipotezy najszybciej (najmniejsze liczebności próbek dla obydwu wartości β) uzyskano w przypadku ziaren żyta dla rzutowego pola powierzchni oraz ziaren owsa dla długości ziaren.

Bibliografia

- Ciećko Z.** 2003. Ocena jakości i przechowalność produktów rolnych. WUWM Olsztyn. ISBN 83-7299-281-9.
- Fisz M.** 1958. Rachunek prawdopodobieństwa i statystyka matematyczna. PWN Warszawa.
- Krysicki W., Bartos J., Dyczka W., Królikowska K., Wasilewski M.** 1998. Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Cz. 2. PWN Warszawa. ISBN 83-01-11384-7.
- Lukaszuk J.** 2005. Wstępna ocena wpływu sposobu formowania złoża ziarna pszenicy na opór przepływu powietrza. Acta Agrophysica. Nr 6(3). s. 709-714.
- Marek T., Noworol C.** 1987. Analiza sekwencyjna w badaniach empirycznych. PWN Warszawa. ISBN 83-01-07274-1.
- Siegmund D.** 1980. Sequential χ^2 and F tests and related confidence intervals. Biometrika 67, 2. s. 389-402.
- Zdunek A., Konsky R., Cybulska J., Konstankiewicz K., Umera M.** 2007. Visual texture analysis for cell size measurements from confocal image. International Agrophysics, 21, 4, s. 409-414.

APPLYING SEQUENTIAL PROCEDURE IN THE ANALYSIS OF GEOMETRICAL CHARACTERISTICS FOR SEEDS OF SELECTED CEREAL PLANTS

Abstract. The scope of the research covered seeds of selected cereal plants (barley, oats, wheat, triticale, rye) with moisture content: 14, 16 and 18%. Basic geometrical characteristics were measured using a computer image analysis system. Sequential tests were employed in the analysis of moisture content impact on projective surface area and grain length.

Key words: corn seeds, computer image analysis, sequential test

Adres do korespondencji

Dorota Domagała; e-mail: dorota.domagala@up.lublin.pl
Katedra Zastosowań Matematyki i Informatyki
Uniwersytet Przyrodniczy w Lublinie
ul. Akademicka 13
20-950 Lublin