

*Andrzej Kornacki, Mirosława Wesołowska-Janczarek
Katedra Zastosowań Matematyki i Informatyki
Uniwersytet Przyrodniczy w Lublinie*

O WERYFIKOWANIU POPRAWNOŚCI MATEMATYCZNYCH MODELI PROCESÓW W OPARCIU O DANE EMPIRYCZNE

Streszczenie

Opisywanie przebiegu procesów modelami matematycznymi jest problemem, który interesuje wielu badaczy w różnych dziedzinach wiedzy. Jednakże ustalenie kształtu modelu wymaga jeszcze zweryfikowania jego poprawności w oparciu o dane empiryczne. Praca zawiera propozycje metod statystycznych, które mogą być wykorzystywane do przeprowadzenia takiej weryfikacji. Zaproponowane metody zostały zilustrowane na przykładach wyników badań uzyskanych podczas eksperymentów z dziedziny inżynierii rolniczej.

Słowa kluczowe: matematyczne modele procesów, weryfikacja modeli, dane empiryczne

Wstęp

W inżynierii systemów, podobnie jak i w innych dziedzinach wiedzy, poznanie przebiegu procesu jest związane z konstrukcją modelu opisującego system i jego działanie. Problemy te są szeroko przedstawione na przykład w książce Powierży [1997]. Ze względu na to, że istnieje duża różnorodność systemów i procesów, jest też możliwość tworzenia różnego typu modeli, a ich systematykę można znaleźć między innymi w książkach Pabisa [1985] czy cytowanej już wcześniej Powierży [1997]. Jak stwierdza Powierża, wśród różnego typu modeli grupa zwana modelami matematycznymi, symbolicznymi lub formalnymi „stanowi najbardziej reprezentatywną grupę” (§ 9.7). W modelach tych wyróżnia się zbiór elementów, czyli zmiennych charakteryzujących system oraz relacje wiążące te elementy. W procesie tworzenia modelu bardzo ważnym etapem jest badanie adekwatności zaproponowanego modelu lub też po prostu jego weryfikacja. Polega ona na ogół na porównaniu wyników liczbowych uzyskanych z modelu z wynikami uzyskanymi z badań empirycznych. Powstaje zatem problem, jakie metody statystyczne można użyć do ich porównania.

Morrison [1990, s. 16] wśród modeli matematycznych wyróżnia trzy grupy: są to modele a) deterministyczne, b) deterministyczne z prostymi wielko-

ściami losowymi (np. błędami pomiarów), c) stochastyczne. W tej pracy zajmować się będziemy także trzema typami modeli matematycznych. Przynależność do grup będzie zależała od tego czy:

- 1^o model został sformułowany na podstawie rozważań teoretycznych lub występujące w nim współczynniki przy zmiennych niezależnych zostały określone na podstawie wcześniej posiadanych danych innych niż te wyniki z eksperymentu, które mają być użyte do weryfikacji modelu.
- 2^o współczynniki w modelu zostały częściowo określone na podstawie teoretycznych rozważań, a częściowo na podstawie danych eksperymentalnych użytych następnie także do weryfikacji modelu.
- 3^o model jest empiryczny, czyli rozważana jest funkcja regresji dopasowana do danych z eksperymentu.

Propozycje wykorzystania metod statystycznych do weryfikacji modeli dla każdego z tych przypadków zostaną przedstawione w dalszej części pracy i zilustrowane na przykładach. Jako przykłady wykorzystano dane dotyczące: modelowania zapotrzebowania na moc efektywną i pracę jednostkową toporowego zespołu rozdrabniającego sieczkarni polowej w oparciu o wyniki przeprowadzonych badań stacjonarnych [Wardecki 2006] oraz badania zmian wysokości brykietów w czasie pod wpływem zastosowanego różnego nacisku.

Warto jeszcze zauważyć, że metody, które można wykorzystać do weryfikowania poprawności modeli z grupy pierwszej mogą też być użyte dla modeli z grupy drugiej, ale nie odwrotnie, gdyż metody statystyczne możliwe do wykorzystania w grupie drugiej wymagają założeń dotyczących rozkładu zmiennej losowej (badanej cechy), co nie jest potrzebne dla modeli z grupy pierwszej, gdzie porównuje się tylko strukturę obu zbiorów danych.

Metody statystyczne do weryfikacji modeli matematycznych z grup pierwszej i drugiej

Spośród wielu znanych metod statystycznych, które mogą być wykorzystane do weryfikowania zaproponowanych modeli matematycznych zostaną wybrane te, które badają zgodność struktury danych, gdy model został wyznaczony na podstawie rozważań teoretycznych lub zgodność rozkładów, gdy niektóre współczynniki w funkcyjnej postaci modelu zostały wyznaczone na podstawie danych empirycznych użytych następnie do jego weryfikacji. W każdym z tych przypadków dysponuje się dwoma zbiorami danych. Pierwszy to zbiór wartości wyznaczonych na podstawie modelu, a drugi to zbiór wartości empirycznych, czyli uzyskanych z pomiaru wielkości badanej cechy aktualnie interesującej eksperymentatora.

Weryfikacja modelu powinna dostarczyć wniosek o zgodności struktury tychże zbiorów w przypadku modeli z grupy pierwszej lub jeśli potraktować pierwszy z rozważanych zbiorów jako wynikający z modelu probabilistycz-

nego [Pabis 1985], czy też w przypadku modelu należącego do grupy drugiej na podstawie posiadanych prób uzyskać wnioski o zgodności rozkładów rozważanych zmiennych losowych.

Dla porównania struktur mogą być wykorzystane następujące wskaźniki:

Wskaźnik determinacji [Makuć, Urbanek 2004]

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - Y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}; \quad R^2 \leq 1 \quad (1)$$

gdzie y_i są wartościami badanej cechy uzyskanymi z pomiarów, Y_i odpowiadającymi im wartościami wyznaczonymi z modelu, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, a n jest liczbą

wartości w każdym ze zbiorów. Im R^2 jest bliższe 1, tym jest lepsza zgodność danych z modelem z danymi empirycznymi. Wartości R^2 podawane są także w procentach.

Współczynnik zbieżności [Makuć, Urbanek 2004]

$$\varphi^2 = \frac{\sum_{i=1}^n (y_i - Y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}; \quad \varphi^2 \geq 0 \quad (2)$$

gdzie użyte we wzorze oznaczenia są takie same, jak dla współczynnika determinacji. Ponadto $R^2 + \varphi^2 = 1$, im φ^2 jest bliższe zeru, tym jest większa zgodność obu zbiorów wartości. Współczynnik ten także może być podawany w procentach.

Względny wskaźnik podobieństwa struktur [Makuć, Urbanek 2004, s. 92]

$$Z = \left(\sum_{i=1}^c \min(w_i) \right) / \left(\sum_{i=1}^c \max(w_i) \right) \quad (3)$$

Dla obliczenia tego wskaźnika, wartości w obu zbiorach muszą być najpierw uporządkowane w szeregi rozdzielcze o takiej samej ilości i długości przedziałów. $w_i = \frac{n_i}{n}$ dla kolejnych przedziałów w szeregach rozdzielczych, c jest liczbą przedziałów, i -kolejnym numerem przedziału, $i = 1, \dots, c$, n_i - liczebnością odpowiadającą i -temu przedziałowi, n -liczebnością każdego ze zbiorów wartości, $\min(w_i)$ to mniejsza z dwóch częstości dla i -tego przedziału, $\max(w_i)$ to większa z nich. (Częstości w_i nazywane są przez ekono-

mistów wskaźnikami struktury). Wartość tego współczynnika zależy od obranej długości przedziałów. Wskaźnik Z przyjmuje wartości z przedziału $\langle 0, 1 \rangle$. Wartość $Z = 1$ wskazuje, że dwie porównywane struktury są identyczne, a gdy $Z = 0$ to struktury są różne. Może być podawany w procentach.

Średni błąd względny

$$W = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - Y_i|}{Y_i} \quad (4)$$

gdzie użyte oznaczenia są takie same jak dla współczynników determinacji i zbieżności. Może on być też wyrażany w procentach. Przyjmuje on wartości z przedziału $\langle 0, 1 \rangle$. Wartość W bliższa zeru świadczy o lepszej zgodności obu zbiorów danych, zaś bliższa jedynki o ich zróżnicowaniu. Może być wyrażana w procentach.

Współczynnik dopasowania

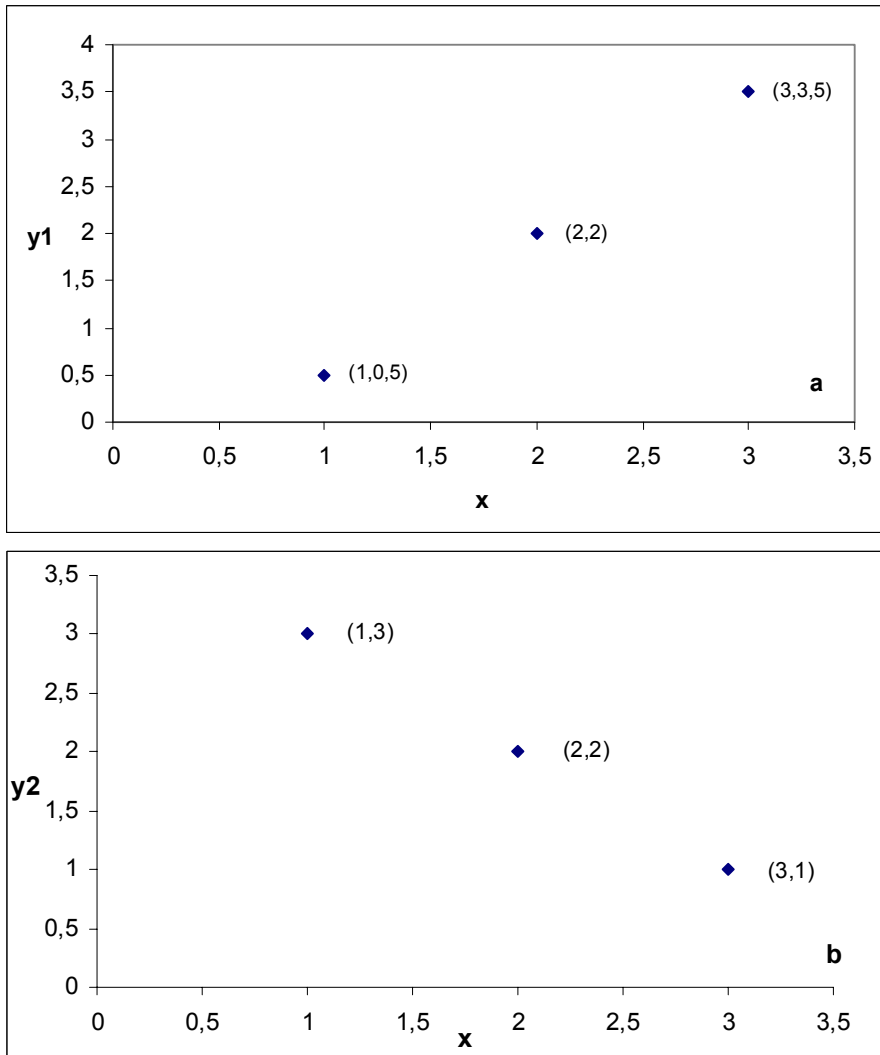
$$d = 1 - \frac{1}{n} \left(\frac{\sum_{i=1}^n |y_i - Y_i|}{\max_i |y_i - Y_i|} \right), \quad 0 \leq d \leq 1 \quad (5)$$

gdzie $\max_i |y_i - Y_i|$ oznacza największą spośród wartości bezwzględnej różnic między odpowiednimi wartościami obserwowanymi i ocenionym z modelu. Wartość d bliska 1 wskazuje na dobre dopasowanie modelu do danych eksperymentalnych, zaś d bliskie 0 – na słabe dopasowanie. Także może być wyrażony w procentach.

Jeśli przyjmie się, że model teoretyczny jest modelem stochastycznym, to wyznaczenie powyższych wskaźników nie jest wystarczające. Powierża [1997] do weryfikacji adekwatności modelu zaleca użycie testu t -Studenta dla porównania średnich lub testu istotności dla współczynnika korelacji. Ponieważ oba testy wymagają spełnienia założenia o rozkładzie normalnym rozważanych cech, a test t -Studenta dla zmiennych niezależnych wymaga jeszcze dodatkowo założenia o równości wariancji problem ten wymaga szerszej dyskusji. Wprawdzie test t -Studenta jest dość odporny na odstępstwa od normalności rozkładów, ale jeżeli nic nie wiadomo o kształcie rozkładów i ich wariancji, to stwierdzenie równości średnich nie wiele mówi o podobieństwie rozkładów rozważanych zmiennych. Wystarczy tu, jako przykład, podać dwa proste zbiory liczbowe wartości y dla ustalonych $x_1 = 1$, $x_2 = 2$, $x_3 = 3$.

Niech $y_{11} = 0,5$, $y_{12} = 2$, $y_{13} = 3,5$ oraz $y_{21} = 3$, $y_{22} = 2$, $y_{23} = 1$.

Łatwo zauważyć, że $\bar{y}_1 = \frac{0,5 + 2 + 3,5}{3} = 2$ i $\bar{y}_2 = \frac{3 + 2 + 1}{2} = 2$, zatem średnie są takie same, ale położenie tych punktów różni się, co widać na rysunku. Również wariancje w obu zbiorach nie są równe.



Rys. 1. Punkty odzwierciedlające zbiór wartości o tej samej średniej a różnym kierunku zmian

Fig. 1. Points showing the set of values of the same mean and different direction of changes

Nie można zatem z warunku równości średnich wnioskować o podobieństwie zbiorów danych, a tym samym i rozkładów. Warto jeszcze zwrócić uwagę na to, że nieodrzućcenie hipotezy zerowej w każdym z testów nie jest równoznaczne z jej przyjęciem. W tym przypadku nie wiemy bowiem bez

dalszych badań jak duży jest błąd II rodzaju, polegający na przyjęciu hipotezy zerowej mimo, że nie jest ona prawdziwa. Trzeba też pamiętać o tym, że zwiększenie mocy testu jest możliwe przez powiększenie liczebności próby. Zatem test t -Studenta poprawnie mógłby być zastosowany jeżeli zostanie poprzedzony testem sprawdzającym normalność obu rozkładów i testem sprawdzającym równość wariancji lub testem dla współczynnika korelacji sprawdzającym $H_0 : \rho_{xy} = 0$ przeciw $H_1 : \rho_{xy} > 0$.

Jeżeli rozważane cechy nie mają rozkładu normalnego, to do porównania ich rozkładów można wykorzystać testy nieparametryczne. Testem sprawdzającym hipotezę o identyczności rozkładów cechy w dwóch niezależnych populacjach jest test λ -Kolmogorowa-Smirnowa [Domański 1990]. Wymaga on dość dużej liczebności prób.

W przypadku, gdy pewne parametry w modelu teoretycznym są szacowane na podstawie danych empirycznych, wykorzystywanych równocześnie do weryfikacji modelu, nie można przyjąć założenia o niezależności wyników w obu zbiorach. Uwzględnienie tego warunku wskazuje, że do weryfikacji modelu należy wybrać testy sprawdzające hipotezę o identyczności rozkładów przy braku założenia o niezależności zmiennych.

W przypadku normalności rozkładów właściwy byłby test t -Studenta dla zmiennych połączonych (dodatnio skorelowanych), a dla dowolnych innych rozkładów mogą być wybrane test znaków [Domański 1990] lub test rangowych znaków, zwany też testem Wilcoxon [Domański 1990]. Oba te testy należą do grupy testów nieparametrycznych.

Weryfikowanie adekwatności empirycznych modeli regresyjnych

Wśród modeli regresyjnych można wyróżnić funkcje regresji liniowe i nieliniowe, ze względu na współczynniki przy zmiennych niezależnych. W grupie liniowych funkcji regresji najczęściej wykorzystuje się regresję wielokrotną (liniową k zmiennych) oraz regresję wielomianową.

Weryfikacja modelu liniowego może obejmować dwa problemy: sprawdzanie „dobroci dopasowania” funkcji regresji do danych empirycznych oraz wybór właściwego modelu z kilku możliwych, jak na przykład odpowiedź na pytanie, którego stopnia wielomian lepiej będzie opisywał zmiany zmiennej zależnej w określonym zbiorze zmiennej niezależnej, czy które zmienne niezależne należy uwzględnić, a które pominąć w przypadku regresji wielokrotnej.

Dobroć dopasowania funkcji regresji do danych najczęściej sprawdza się obliczając współczynnik determinacji według wzoru (1) lub według wzoru (6) określającego współczynnik R^2 jako kwadrat współczynnika korelacji między wartościami empirycznymi y_i , a odpowiadającymi im wartościami Y_i wyliczonymi na podstawie funkcji regresji.

$$R^2 = \frac{\frac{1}{n} \sum_{i=1}^n y_i Y_i - \bar{y} \cdot \bar{Y}}{\left(\frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{y}^2 \right) \left(\frac{1}{n} \sum_{i=1}^n Y_i^2 - \bar{Y}^2 \right)} = r_{yY}^2 \quad \text{przy } 0 < R^2 < 1 \quad (6)$$

Należy jednak pamiętać, że gdy model jest nieliniowy, np. logistyczny lub źle wybrany jest kształt funkcji regresji, współczynnik R^2 może mieć wartość ujemną. Warto również zwrócić uwagę na to, że przy dopasowywaniu liniowej regresji wielokrotnej (zwanej też regresją wieloraką) zamiast używać zwykłego współczynnika determinacji R^2 polecany jest tak zwany skorygowany (lub poprawiony) współczynnik determinacji \bar{R}^2 w postaci

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-k-1} \quad (7)$$

gdzie k jest liczbą zmiennych niezależnych.

Ma on tę zaletę, że nie zależy od liczby zmiennych niezależnych. Ponadto, jak podają Józwiak i Podgórski [2006], poprawiony współczynnik \bar{R}^2 może być więc wykorzystany do ustalenia, czy wprowadzone do modelu nowe zmienne poprawiają stopień opisu przebiegu zmian badanej cechy. Wartość współczynnika \bar{R}^2 zmniejsza się, jeśli wprowadzone są nowe zmienne, które nie powodują znaczącego powiększenia się sumy kwadratów odchyłeń wyjaśnionej przez regresję, podczas gdy wartość zwykłego współczynnika determinacji wzrasta wraz ze wzrostem liczby zmiennych niezależnych.

Do sprawdzania, czy model został wybrany właściwie, lub inaczej, czy regresja jest istotna, wykorzystuje się następujące testy statystyczne: test t -Studenta do sprawdzania istotności dla poszczególnych współczynników (w modelu regresji liniowej: wielomianowej lub wielokrotnej) oraz test analizy wariancji F dla pełnej funkcji regresji. Do wyboru właściwego (najlepszego) spośród kilku modeli, które mogłyby być wzięte pod uwagę jako funkcje regresji, w konkretnej sytuacji, można poza współczynnikiem R^2 użyć też współczynnika C_p podanego przez Mallowsa [1973]. Współczynnik ten często używany w badaniach ekonomicznych wyrażany jest wzorem:

$$C_p = \frac{RSS}{s^2} - (n - 2p) \quad (8)$$

gdzie $RSS = \sum_{i=1}^n (y_i - Y_i)^2$ jest resztową sumą kwadratów odchyłeń, $s^2 = \frac{nS_y^2}{n-1}$ jest estymatorem wariancji cechy y , n – liczbą obserwacji, a p – liczbą szacowanych parametrów.

Wartość $C_p = p$ lub bliska p wskazuje na dobre dopasowanie modelu, natomiast im $C_p > p$, tym dopasowanie jest gorsze. Wskaźnik ten można stosować przy eliminowaniu pewnych zmiennych z modelu w regresji wielokrotnej. Oprócz wskaźników R^2 , (\bar{R}^2) i C_p można praktycznie wykorzystywać też i inne kryteria wyboru modeli, które są opisane na przykład w książce Maddali [2006] czy też Daniela i Wooda [1980] oraz stosować wskaźniki oceniające jakość modelu oparte na analizie reszt, które można znaleźć w książce Dobosza [2001].

Przykłady zastosowań

Przykład 1

Wardecki [2006] na podstawie przeprowadzonych badań stanowiskowych opracował modele matematyczne, opisujące zapotrzebowanie na moc efektywną i pracę jednostkową toporowego zespołu sieczkarni polowej przy uwzględnieniu wpływu dodatkowych elementów wspomagających rozdrabnianie roślin, takich jak listwy promieniowe, łopatki rozrzutnika, płytka denna czy szczelina robocza. Modele te mają następującą postać:

- dla mocy efektywnej

$$P_t = L_j \frac{bh}{2\pi} \frac{z}{\pi} + \lambda \frac{(1+k)^2 \cos^2 \tau + f^2 \sin^2 \tau}{2} q_m^2 r^2 + \mu k_r f_p q_m^2 r_t^2 (\beta_t - \beta_{td}) + \mu_d k_r f_p q_m^2 r_t^2 i_k \frac{b_k}{\cos \alpha_k} + \mu_s q_m^2 r_r l_d \frac{s_m - s}{s_m} + \lambda_h \frac{q_m^2 r_t^2}{2} \quad (9)$$

- dla pracy jednostkowej

$$L_{tsm} = L_j \frac{bh}{2\pi q_m (1-w_r)} \frac{z}{\pi} + \lambda \frac{(1+k)^2 \cos^2 \tau + f^2 \sin^2 \tau}{2(1-w_r)} r^2 + \mu \frac{k_r f_p r_t^2 (\beta_t - \beta_{td})}{1-w_r} + \mu_d \frac{k_r f_p r_t^2 i_k b_k}{(1-w_r) \cos \alpha_k} + \mu_s \frac{r_t^2 l_d (s_m - s)}{(1-w_r) s_m} + \lambda_h \frac{r_t^2}{2(1-w_r)} \quad (10)$$

W modelach tych wielkości nieznanymi współczynnikami L_j , λ , μ , μ_d , μ_s oraz λ_k zostały oszacowane na podstawie danych eksperymentalnych przy użyciu programów zawartych w pakiecie Statgraphics v.4.1.

Modele te można zaliczyć do grupy II ze względu na to, że część współczynników wyznaczono na podstawie rozważań teoretycznych, a część na podstawie wyników pomiarów. Wykorzystując dwa zbiory wartości: obliczonych zgodnie z podanymi modelami oraz wartości empirycznych obliczono kolejne wskaźniki i zastosowano odpowiednie testy, które rozważono w poprzedniej części pracy:

1. Współczynnik determinacji R^2 obliczony według wzoru (1)

dla mocy efektywnej P : $R_P^2 = 83\%$

dla pracy jednostkowej L : $R_L^2 = 55\%$

2. Współczynnik zbieżności φ^2 obliczony według wzoru (2)

dla mocy efektywnej P : $\varphi_P^2 = 17\%$

dla pracy jednostkowej L : $\varphi_L^2 = 45\%$

Oba współczynniki wskazują, że lepiej jest dopasowany model mocy efektywnej niż pracy jednostkowej.

3. Względny wskaźnik podobieństwa struktur (wzór (3))

dla mocy efektywnej P : $Z_P = 0,752$ lub $Z_P = 75\%$

dla pracy jednostkowej L : $Z_L = 0,403$ lub $Z_L = 40\%$

Ten współczynnik także potwierdza wniosek o lepszym dopasowaniu modelu dla mocy efektywnej.

4. Średni błąd względny W (wzór (4)):

dla mocy efektywnej P : $W_P = 11\%$

dla pracy jednostkowej L : $W_L = 10\%$

Wartości obu ostatnich współczynników dla rozważanych modeli wykazują niewielkie różnicowanie, a mniejszy średni błąd względny sugeruje jako jedyny ze wskaźników lepsze dopasowanie modelu pracy jednostkowej niż dla mocy efektywnej.

Jeżeli potraktuje się oba zbiory wartości jako próby losowe, można badać zgodność rozkładów zmiennych losowych reprezentowanych przez te próby. Ponieważ rozważane zmienne nie są niezależne zostaną wykorzystane testy znaków i Wilcoxa zarówno dla zbadania zgodności rozkładów mocy efektywnej (P), jak i pracy jednostkowej (L). Przypomnijmy, że test znaków sprawdza hipotezę zerową o równości dystrybuant dla zmiennych z dwóch populacji $H_0 : F_1(x) = F_2(y)$ przeciw $H_1 : F_1(x) \neq F_2(y)$. Obie próby muszą mieć tę samą liczebność n . Obliczamy różnice $(x_i - y_i)$ dla $i = 1, \dots, n$ i zliczamy ile jest różnic dodatnich, oznaczamy tę liczbę r^+ oraz ile jest różnic ujemnych (r^-). Przez r oznaczamy mniejszą z wartości r^+ i r^- , czyli:

$$r = \min(r^+, r^-) \quad (11)$$

Statystyka r jest tu funkcją testową, która przy prawdziwej H_0 ma rozkład dwumianowy. Lewostronny obszar krytyczny określony jest równością $P(r \leq r_\alpha) = \alpha$ i r_α odczytuje się z odpowiednich tablic (patrz np. [Domański 1990 s. 168]).

Wracając do przykładu mamy:

- dla mocy efektywnej P :

$r_P^+ = 43$, $r_P^- = 27$; $r_P = \min(r_P^+, r_P^-) = 27$, przy $\alpha = 0,01$; $r_{0,01} = 24$ zatem $r_P > r_\alpha$ i nie ma podstaw do odrzucenia H_0 , co potwierdza zgodność rozkładów czyli można przyjąć, że model (4.1) dobrze opisuje moc efektywną,

- dla pracy jednostkowej L :

$r_L^+ = 43$, $r_L^- = 27$; $r_L = \min(r_L^+, r_L^-) = 27$, przy $\alpha = 0,01$; $r_{0,01} = 24$; mamy więc $r_L > r_\alpha$ i też nie ma podstaw do odrzucenia H_0 , można więc przyjąć, że model (4.2) dobrze opisuje pracę jednostkową.

Zobaczmy jeszcze, jakie wnioski uzyskamy stosując test Wilcoxona. Dla zastosowania tego testu, tak jak w poprzednim teście, należy obliczyć wartości różnic $x_i - y_i$, ale do dalszych działań należy brać pod uwagę ich wartości bezwzględne $|x_i - y_i|$. Tym wartościom bezwzględnym należy przypisać rangi, a następnie obliczyć osobno sumę rang różnic, które były dodatnie oznaczając ją T^+ oraz sumę rang dla różnic ujemnych T^- . Funkcja testowa jest równa mniejszej wartości z tych sum, czyli

$$T = \min(T^+, T^-) \tag{12}$$

Obszar krytyczny testu wyznacza relacja: $P(T \leq T_\alpha) = \alpha$. Wartości krytyczne można odczytać z tablic zamieszczonych na przykład w książce Domańskiego [1990, s. 269]. Jeżeli liczebność prób jest duża ($n > 25$), to przy prawdziwej H_0 o równości dystrybuant można wykorzystać inną postać funkcji testowej, która ma rozkład asymptotycznie normalny $N(0,1)$. Ma ona postać:

$$u^o = \frac{T - E(T)}{D(T)}, \tag{13}$$

gdzie $E(T) = \frac{1}{4}n(n+1)$ i $D^2(T) = \frac{1}{24}n(n+1)(2n+1)$, a $D(T) = \sqrt{D^2(T)}$.

Jeśli $|u^o| > u_\alpha$ to H_0 odrzucamy tak jak przy innych testach, gdzie funkcja testowa ma rozkład normalny.

Uwzględniając wartości liczbowe z przykładu, gdzie $n = 70$:

- dla mocy efektywnej P :

$E_P(T) = 1242,5$; $D_P^2(T) = 29198,8$ i $D_P(Y) = 170,876$, $T_P^+ = 1512$; $T_P^- = 973$
zatem $T_P = 973$

$$u_P^0 = \frac{973 - 1242,5}{170,876} = -1,58$$

przy $\alpha = 0,01$, $u_\alpha = 2,576$ zatem $u^0 < -u_\alpha$ czyli nie odrzucamy H_0 , a zatem można przyjąć, że model dobrze opisuje moc efektywną,

- dla pracy jednostkowej L :

$E_L(T) = 1242,5$; $D_L^2(T) = 29198,8$ i $D_L(T) = 170,876$, $T_L^+ = 1654$; $T_L^- = 831$
więc $T_L = 831$

$$u_L^0 = \frac{831 - 1242,5}{170,876} = -2,41$$

przy $\alpha = 0,01$, $u_\alpha = 2,576$ czyli $|u_L^0| < u_\alpha$ i podobnie jak poprzednio nie odrzucamy H_0 , czyli przyjmujemy iż dane empiryczne potwierdzają, że przedstawiony model dobrze opisuje pracę jednostkową.

Porównajmy jeszcze otrzymane w wyniku zastosowania tego testu wnioski z wnioskami uzyskanymi po zastosowaniu testu większej mocy t -Studenta dla zmiennych połączonych. Użyty wcześniej test normalności λ -Kolmogorowa dla obu cech nie odrzucił sprawdzanej hipotezy (dla danych empirycznych dotyczących mocy efektywnej $\lambda_e^0 = 0,613$ przy $\lambda_{0,05} = 1,358$; a dla danych obliczanych z modelu $\lambda_m^0 = 0,485$; natomiast dla pracy jednostkowej $\lambda_e^0 = 0,819$ i $\lambda_m^0 = 0,280$ przy $\lambda_{0,05}$ takim samym, jak dla mocy). Stosując znany test t -Studenta dla zmiennych połączonych dla obu cech, to jest mocy efektywnej i pracy jednostkowej, otrzymujemy:

- dla mocy efektywnej: $t^0 = 1,565$ przy $t_{0,05;69} = 1,995$, co pozwala przyjąć hipotezę zerową i stwierdzić, że model poprawnie opisuje zmiany badanej cechy,

- dla pracy jednostkowej: $t^0 = 2,947$ przy $t_{0,05;69} = 1,995$ i $t_{0,01} = 2,649$, a ponieważ nawet $t^0 > t_{0,01}$, zatem hipotezę o równości średnich danych empirycznych i obliczonych z modelu należy odrzucić, dlatego nie wydaje się, aby przyjęty model dobrze opisywał zmiany badanej cechy.

Zauważmy jeszcze, że wśród wcześniej podanych obliczonych wskaźników, ich wartości obliczone dla pracy jednostkowej także nie potwierdzały wysokiej zgodności danych empirycznych z modelowymi.

Przykład 2

W badaniach dotyczących pewnych własności brykietów wytwarzanych z resztek poźniwnych kukurydzy, przeprowadzonych w Katedrze Maszynoznawstwa Rolniczego Akademii Rolniczej w Lublinie, obserwowano ich wysokość y w zależnościach od czasu po sprasowaniu przy stosowaniu różnego nacisku. Uwzględniono pomiary wysokości bezpośrednio po sprasowaniu, oraz w 15, 30, 45 i 60 minut po nim. Używając funkcji regresji wielomianowej próbowano ustalić jak kształtuje się wysokość brykietu jako funkcja czasu. Na podstawie dwóch zestawów zamieszczonych w tabelce poniżej przy uwzględnionym nacisku 10 t i 15 t dokonano odpowiednich pomiarów.

Tabela 1. Wysokość brykietów w różnym czasie po sprasowaniu
Table 1. The height of briquetts at different time after pressing

x- czas (w min)	y- wysokość brykietów (mm)	
	zestaw I	zestaw II
po sprasowaniu	40,35	37,24
15	43,51	39,82
30	42,85	41,22
45	43,07	42,61
60	43,19	42,80

Do modelowania wybrano wielomiany stopnia pierwszego, drugiego i trzeciego. Dla każdego z nich policzono współczynniki R^2 oraz C_p , według wzorów (6) i (8), na podstawie których ustalono stopień wielomianu najlepiej opisujący relację między wysokością a czasem. Dla zestawu I obserwacji przy nacisku 10 ton wyznaczone równania regresji i odpowiednie współczynniki R^2 i C_p , były następujące:

- wielomian stopnia pierwszego

$$y = 0,0349x + 41,546; \quad R^2 = 0,421 \quad \text{i} \quad C_p = 1,32 ,$$

- wielomian stopnia drugiego

$$y = -0,0017x^2 + 0,134x + 40,803; \quad R^2 = 0,7171 \quad \text{i} \quad C_p = 2,1577 ,$$

- wielomian stopnia trzeciego

$$y = 0,5 \cdot 10^{-9}x^3 + 0,0099x^2 + 0,3117x + 40,431; \quad R^2 = 0,9293 \quad \text{i} \quad C_p = 2403,9 .$$

Spośród przytoczonych powyżej modeli regresyjnych należy wybrać model stopnia drugiego. Wskazuje na to wyraźnie wartość C_p najbliższa $p = 3$. Wiadać też, że R^2 nie jest tu najlepszą miarą dopasowania, ponieważ wartość jego wzrasta wraz ze wzrostem stopnia wielomianu. Obliczone wartości tych wskaźników dla drugiego zestawu danych przy nacisku 15 ton są następujące:

- regresja liniowa

$$y = 0,0927x + 37,956; \quad R^2 = 0,9184 \quad \text{ i } \quad C_p = -0,67$$

- regresja kwadratowa

$$y = -0,015x^2 + 0,184x + 37,272; \quad R^2 = 0,9962 \quad \text{ i } \quad C_p = 1,0167$$

- regresja kubiczna

$$y = -0,7 \cdot 10^{-5}x^3 + 0,0015x^2 + 0,183x + 37,274; \quad R^2 = 0,9962 \quad \text{ i } \quad C_p = 42,105$$

Tym razem także najbliższy wartości $p = 3$ jest $C_p = 1,0167$, zatem należy przyjąć model regresji kwadratowej.

Dla pełniejszego porównania, czy przebieg zmian wysokości w czasie w badanych zestawach różniących się wielkością zastosowanego nacisku jest podobny a więc, czy można byłoby opisać go jednym wspólnym równaniem należałoby zastosować odpowiedni test. Porównanie równań regresji dokonywane jest z uwzględnieniem jednakowych lub różnych wariancji cechy y w badanych zbiorach wartości. Odpowiednie testy służące do takiego porównania można znaleźć w książce Sebera [1977] i licznych pracach, jak np. Chow [1960], czy Szczepanik i Wesołowska-Janczarek [2006]. Bardziej zaawansowanymi wielozmiennymi metodami, uwzględniającymi porównanie dowolnie wielu równań regresji wielomianowych dla skorelowanych obserwacji wyrażonych jako funkcja czasu, są metody krzywych wzrostu, których przegląd zawarty jest w pracy Wesołowskiej-Janczarek [1993].

Bibliografia

Chow G.C. 1960. Test of equality between sets of coefficients in two linear regressions

Daniel C., Wood F.S. 1980. Fitting Equations to Data, Computer Analysis of Multifactor Data. Sec. ed., John Wiley and Sons, New York

Dobosz M. 2001. Wspomagana komputerowo statystyczna analiza wyników badań, Akademicka Oficyna Wydawnicza EXIT, Warszawa

Domański Cz. 1990. Testy statystyczne, PWE, Warszawa

Jóźwiak J., Podgórski J. 2006. Statystyka od podstaw. Wyd. VI zmienione, Państwowe Wydawnictwo Ekonomiczne, Warszawa

- Maddala G.S. 2006: Ekonometria. PWN, Warszawa
- Makuć W., Urbanek-Krzysztofiak D. 2004. Metody opisu statystycznego, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk
- Mallows C.L. 1973. Some Comments on Cp. *Technometrics*, 15: 661-675
- Morrison D.F. 1990. Wielowymiarowa analiza statystyczna. PWN, Warszawa
- Pabis St. 1985. Metodologia i metody nauk empirycznych. PWN, Warszawa
- Powierża L. 1997. Elementy inżynierii systemów. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa
- Seber G.A.F. 1977. *Linear Regression Analysis*. J. Wiley & Sons, New York
- Szczepanik M., Wesołowska-Janczarek M. 2006. Porównanie testów do sprawdzania identyczności dwóch modeli regresji w przypadku heteroscedastyczności, *Coll. Biometr.*, 36: 257-266
- Wardecki P. 2006. Wpływ parametrów technicznych i eksploatacyjnych na obciążenia energetyczne toporowego zespołu rozdrabniającego. Praca doktorska. SGGW, Warszawa
- Wesołowska-Janczarek M. 1993. Krzywe wzrostu – przegląd metod i nowe problemy. *Listy Biometryczne*, 30(2): 73-96