

**Czesław Domański**

Uniwersytet Łódzki

---

## WŁASNOŚCI TESTU DAVIDA–HELLWIGA

---

**Streszczenie:** W literaturze statystycznej obserwujemy stały wzrost zainteresowania metodami nieparametrycznymi. Metody te na ogół są oparte na statystykach liczących rangowych, licząco-rangowych, pozycyjnych oraz na liczbie i długości serii. W pracy tej przedstawiony zostanie nie mniej znany test oparty na liczbie pustych cel. Rozważany będzie test zgodności Davida–Hellwiga. Przedstawiona zostanie empiryczna moc tego testu w porównaniu z klasycznymi testami: Kołmogorowa–Lillieforsa oraz Shapiro–Willka dla hipotezy o normalności rozkładu.

**Słowa kluczowe:** testy: Kołmogorowa–Lillieforsa, Shapiro–Willka, Davida–Hellwiga empiryczna moc testów normalności.

### 1. Uwagi wstępne

W literaturze statystycznej obserwuje się stały wzrost zainteresowania metodami nieparametrycznymi. Metody te najczęściej są oparte na statystykach liczących (testy znaków), rangowych, licząco-rangowych, pozycyjnych, opartych na liczbie lub długości serii, na liczbie pustych cel.

Wydaje się, że wzrost zainteresowania tą klasą testów wynika głównie z tego, że podejście nieparametryczne jest bardziej realistyczne, uniwersalne i pewne. Należy jednak podkreślić, że zawsze ważne jest każdorazowe analizowanie zachowania się procedur testowych w niestandardowych problemach badawczych. Istotne są sformułowanie pytania, na które chcemy uzyskać odpowiedź w wyniku zastosowania określonej procedury statystycznej, i wstępne postulaty, które muszą być spełnione, aby nasza odpowiedź była zasadna. Istotnym błędem powodującym dochodzenie do nieprawdziwych wniosków jest stosowanie procedur statystycznych, które są oparte na pewnych, przyjętych z góry założeniach nie dających się zweryfikować na podstawie uzyskanych obserwacji, a także logicznie uzasadnić. Widać więc, że metody statystyczne powinny być traktowane nie tylko jako procedury szczególnego modelu, również lecz także jako procedury uzupełniające przy różnych modelach uwzględniających metody wnioskowania statystycznego.

Uwzględniając różnorodność trudności i niebezpieczeństwa wynikające ze stosowania dość powszechnie tych procedur, statystycy dążą do opracowywania takich

metod, które zależą od jak najmniejszego zestawu założeń. Zwykle takie metody należą do procedur wnioskowania nieparametrycznego zwanych również metodami niezależnymi od postaci rozkładu badanych zmiennych losowych.

W pracy przedstawione zostaną wyniki analizy własności testu Davida–Hellwiga dla weryfikacji hipotezy o normalności rozkładu, którego funkcją testową jest liczba pustych cel.

Testy oparte na liczbie pustych cel dla jednej próby przedstawili David [2] i Okamoto [10]. Następnie teoria pustych cel była rozwijana przez Csorgo i Guttmana [1], Hellwiga [7], Kołczina, Sewastianowa i Czistiakowa [9], Domańskiego i Tomaszewicza [5; 6]. Hellwig [7] zaproponował zastosowanie liczby pustych cel jako statystyki testu normalności dla hipotezy złożonej (tj. gdy parametry statystycznego rozkładu nie są znane), (por. także [8]).

## 2. Rozkład liczby pustych cel

Rozważmy próbę złożoną z  $n$  niezależnych obserwacji wylosowanych z populacji o ciągłej dystrybucji  $F$ :

$$x_1, x_2, \dots, x_n \quad (1)$$

i chcemy zweryfikować hipotezę

$$H_0 : F = F_0 \text{ wobec } H_1 : F \neq F_0, \quad (2)$$

gdzie:  $F_0$  jest znane.

Podzielimy zbiór liczb rzeczywistych na  $m$  rozłącznych podzbiorów (cel), tak aby prawdopodobieństwo tego, że pojedyncza obserwacja wylosowana z określającego hipotezę zerową (1) rozkładu należy do określonej celi wynosiła  $1/m$ .

Rozkład liczby pustych cel  $K$ , tj. tych podzbiorów, które nie zawierają żadnej obserwacji z próby, przy założeniu prawdziwości hipotezy  $H_0$  ma postać:

$$P(K = k) = \binom{m}{k} \sum_{r=0}^{m-k-1} (-1)^r \binom{m-k}{r} \left(1 - \frac{k+r}{m}\right)^n. \quad (3)$$

Wartość oczekiwana i wariancja rozkładu  $K$  odpowiednio dane są wzorami:

$$E(K) = m \left(1 - \frac{1}{m}\right)^n, \quad (4)$$

$$D^2(K) = m(m-1) \left(1 - \frac{2}{m}\right)^n + m \left(1 - \frac{1}{m}\right) n - m^2 \left(1 - \frac{1}{m}\right)^{2n}. \quad (5)$$

Warto zauważyć pewną niedogodność rachunkową stosowania wzoru (3). Wartości bezwzględne składników prawej strony (3) są nieporównywalnie większe od ich sumy i dysproporcja ta szybko rośnie wraz ze wzrostem  $n$ . W związku z tym

pojawiają się trudności związane z zabezpieczeniem należytej dokładności obliczeń. Do rozważań posłużę się następującą modyfikacją wzoru (3):

$$P(K = k) = \binom{m}{k} \sum_{r=0}^{m-k-2} (-1)^r \binom{m-k-2}{r} \times \left[ \left(1 - \frac{k+r+2}{m}\right)^n - 2 \left(1 - \frac{k+r-1}{m}\right)^n + \left(1 - \frac{k+r}{m}\right)^n \right]. \quad (6)$$

Modyfikacja ta zapewniła jednakże dostateczną dokładność dla  $m = n < 50$  (por. [5]).

### 3. Test normalności Davida–Hellwiga

Przedmiotem badań jest test Davida–Hellwiga dla hipotezy o normalności rozkładu z nieznanymi parametrami. W takim przypadku (por. [7]) cele są określone na bazie rozkładu normalnego  $N(\bar{x}, s)$ , gdzie  $\bar{x}$  jest średnią arytmetyczną,  $s$  – odchyleniem standardowym z próby.

Weźmy pod uwagę próbę złożoną z  $n$  niezależnych realizacji  $x_1, x_2, \dots, x_n$  pewnej zmiennej losowej o dystrybuancie  $F$ . Zgodnie z ogólnie akceptowaną symboliką, wprowadźmy wskaźnik w nawiasie jako numer odpowiedniego elementu w próbie uporządkowanej

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}. \quad (7)$$

Zdefiniujemy jeszcze „standaryzowane” elementy próby

$$z_i = \frac{x_i - \bar{x}}{s}, \quad z_i^* = \frac{x_i - \bar{x}}{s^*}, \quad (8)$$

gdzie:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (9)$$

$$s^* = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (10)$$

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}. \quad (11)$$

Rozważmy statystykę Davida–Hellwiga, dla takiego rozkładu.

Numer celi zawierającej standaryzowany element próby  $z_i$  określamy następująco:

$$l_i = \text{entier}(mz_i) + 1 \quad \text{dla } i = 1, \dots, n. \quad (12)$$

Liczba elementów próby należących do  $j$ -ej celi wynosi:

$$m_j = \text{card}\{i: l_i = j\} \quad \text{dla } j = 1, 2, \dots, m. \quad (13)$$

Liczba pustych cel określona wzorem

$$K^* = \text{card}\{j: m_j = 0\} \quad (14)$$

jest sprawdzianem testu Davida–Hellwiga.

Zauważmy, że statystyka (14) jest funkcją zmiennych  $z_{(j)}$ , które tylko w przybliżeniu są niezależne i normalne. Zatem można się spodziewać, że rozkład  $K^*$  różni się od rozkładu  $K$  określonego wzorem (3) przy założeniu, że obserwacje w próbie są generowane niezależnie ze znanego rozkładu.

#### 4. Badanie mocy testu Davida–Hellwiga

Moc testu Davida–Hellwiga porównamy z mocą testu Kołmogorowa–Lillieforsa oraz Shapiro–Wilka.

Test Kołmogorowa–Lillieforsa oparty jest na statystyce postaci:

$$D = \sup \left| \hat{F}(x) - F(x) \right|, \quad (15)$$

gdzie  $\hat{F}$  jest dystrybuantą empiryczną określoną wzorem:

$$\hat{F}(x) = \begin{cases} 0 & \text{dla } x \leq x_1, \\ \frac{1}{n} & \text{dla } x_{(i)} < x \leq x_{(i+1)} \quad (i = 1, 2, \dots, n-1), \\ 1 & \text{dla } x_{(n)}. \end{cases} \quad (16)$$

Statystyka Shapiro–Wilka dana jest w postaci ilorazu form kwadratowych:

$$W = \frac{\left( \sum_{i=1}^r a_i (x_{(n-i+1)} - x_{(i)}) \right)^2}{\sum_{i=1}^r (x_i - \bar{x})^2}, \quad (17)$$

gdzie  $r = \text{entier}(n/2)$ , współczynniki  $a_i$  zaś są tak dobrane, aby wariancja resztowa liniowej regresji między elementami  $x_{(i)}$  a elementami uporządkowanej próby pochodzącej ze standaryzowanego rozkładu normalnego była najmniejsza.

Tablice współczynników  $a_i$  (znormalizowanych, tj. spełniających warunek  $\sum_i a_i^2 = 1$ ) dla  $n \leq 50$  opublikowane zostały w pracy [11], (por. także [4]).

W celu oceny mocy testu Davida–Hellwiga rozważaliśmy cztery rozkłady alternatywne:

(a) rozkład równomierny o gęstości:

$$f_2(x) = \begin{cases} 1 & \text{gdy } 0 \leq x \leq 1, \\ 0 & \text{gdy } x < 0 \text{ lub } x > 1, \end{cases} \quad (18)$$

(b) rozkład „trójkątny symetryczny” o gęstości:

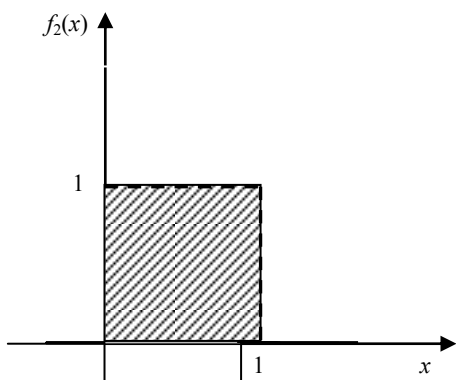
$$f_3(x) = \begin{cases} 4x & \text{gdy } 0 \leq x \leq \frac{1}{2}, \\ 4 - 4x & \text{gdy } \frac{1}{2} < x \leq 1, \end{cases} \quad (19)$$

(c) rozkład „trójkątny asymetryczny” o gęstości:

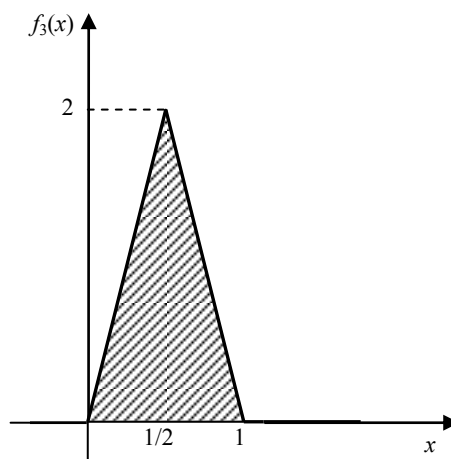
$$f_4(x) = \begin{cases} 2x & \text{gdy } 0 \leq x \leq 1, \\ 0 & \text{gdy } x < 0 \text{ lub } x > 1, \end{cases} \quad (20)$$

(d) rozkład „dwumodalny” o gęstości:

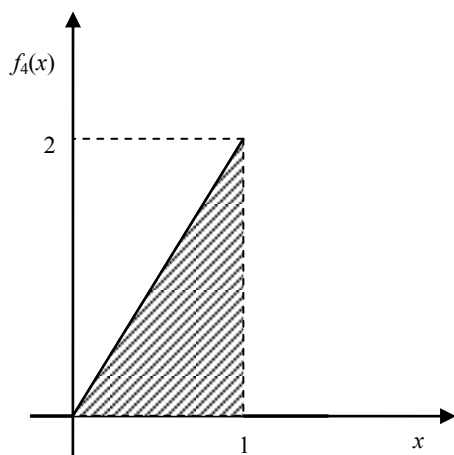
$$f_5(x) = \begin{cases} 2 - 4x & \text{dla } 0 \leq x \leq \frac{1}{2}, \\ 4x - 2 & \text{dla } \frac{1}{2} < x \leq 1, \\ 0 & \text{gdy } x < 0 \text{ lub } x > 1. \end{cases} \quad (21)$$



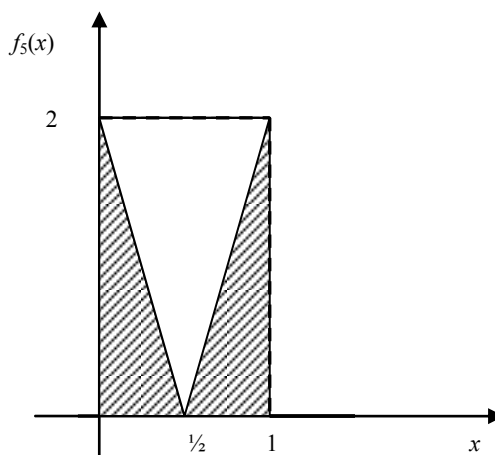
Rys. 1. Rozkład równomierny



Rys. 2. Rozkład trójkątny symetryczny



Rys. 3. Rozkład trójkątny asymetryczny



Rys. 4. Rozkład dwumodalny

Odpowiednie dystrybuanty oznaczymy przez  $F_2, F_3, F_4, F_5$ , zachowując symbol  $F_1$  dla dystrybuanty rozkładu normalnego. Wykresy rozkładów alternatywnych przedstawione są na rys. 1-4.

## 5. Wyniki badania

Badanie przeprowadzono metodą Monte Carlo, na podstawie 5000 repetycji. Dla próby o liczebności  $n = 10, 20, 30, 40, 50$  i poziomu istotności  $\alpha = 0,01; 0,05; 0,10$  porównano moc rozważanych testów, uwzględniając przedstawione cztery rozkłady  $F_\alpha = F_2, F_3, F_4, F_5$ , szacując następujące trzy prawdopodobieństwa:

$$\begin{aligned} \nu_K &= P\{K \geq k_\alpha | F = F_\alpha\}, \\ \nu_W &= P\{W \leq w_\alpha | F = F_\alpha\}, \\ \nu_D &= P\{D \geq d_\alpha | F = F_\alpha\}. \end{aligned} \tag{22}$$

Wartości krytyczne  $k_\alpha, w_\alpha$  oraz  $d_\alpha$  spełniają następujące relacje:

$$\begin{aligned} P\{K \geq k_\alpha | F = F_1\} &\leq \alpha \langle P\{K \geq k_\alpha - 1 | F = F_1\} \rangle, \\ P\{W \leq w_\alpha | F = F_1\} &= \alpha, \\ P\{D \geq d_\alpha | F = F_1\} &= \alpha \end{aligned} \tag{23}$$

i podane są w tab. 1.

**Tabela 1.** Wartości krytyczne dla testów Shapiro–Wilka ( $w_\alpha$ ), Kołmogorowa ( $d_\alpha$ ), Hellwiga ( $k_\alpha$ ) i współczynniki randanizacyjne ( $P_R$ ), (por. wzór (23)).

$\alpha$	$n$	Wartości krytyczne			Współczynniki $P_R$
		$w_\alpha$	$d_\alpha$	$k_\alpha$	
.01	10	.781	.294	6	.053
	20	.868	.231	11	.327
	30	.900	.187	16	.439
	40	.919	.163	20	.438
	50	.930	.146	24	.344
.05	10	.842	.258	6	.069
	20	.905	.190	10	.312
	30	.927	.161	15	.683
	40	.940	.140	18	.039
	50	.947	.125	22	.030
.10	10	.869	.239	5	.121
	20	.920	.174	10	.893
	30	.939	.144	14	.326
	40	.949	.127	18	.649
	50	.955	.114	22	.778

Źródło: na podstawie tablic zamieszczonych w pracy [4].

Ze względu na to, że rozkład statystyki  $K$  jest dyskretny, nie można dobrać wartości  $k_\alpha$  tak, aby  $P\{K \geq k_\alpha\} = \alpha\}$ . W związku z tym szacowaliśmy również moc zranominizowanego wariantu testu Davida–Hellwiga

$$\nu_{KR} = P\{K \geq k_\alpha | F = F_\alpha\} + P_R P\{K = k_\alpha - 1 | F = F_\alpha\}, \quad (24)$$

gdzie

$$P_R = \frac{\alpha - P\{K \geq k_\alpha | F = F_1\}}{P\{K = k_\alpha - 1 | F = F_1\}}. \quad (25)$$

**Tabela 2.** Empiryczna moc badanych testów dla rozkładu równomiernego (%)

$\alpha$	$n$	Test			
		Davida–Hellwiga ( $\hat{\nu}_K$ )	Davida–Hellwiga ( $\hat{\nu}_{KR}$ )	Shapiro–Wilka ( $\hat{\nu}_w$ )	Kołmogorowa– Lillieforsa ( $\hat{\nu}_D$ )
1	2	3	4	5	6
.01	10	15	21	9	21
	20	21	46	36	12
	30	36	75	80	36
	40	46	76	281	55
	50	76	105	436	64
.05	10	15	97	75	84
	20	87	147	188	105

1	2	3	4	5	6
.10	30	124	237	333	132
	40	241	248	686	179
	50	297	303	821	266
	10	134	174	173	144
	20	97	240	340	198
	30	289	358	530	256
	40	241	355	847	332
	50	297	447	933	418

Źródło: obliczenia własne.

**Tabela 3.** Empiryczna moc badanych testów dla symetrycznego rozkładu trójkątnego w (%)

$\alpha$	$n$	Test			
		Davida–Hellwiga ( $\hat{\nu}_K$ )	Davida–Hellwiga ( $\hat{\nu}_{KR}$ )	Shapiro–Wilka ( $\hat{\nu}_w$ )	Kolmogorowa– Lillieforsa ( $\hat{\nu}_D$ )
.01	40	4	10	9	7
	50	4	11	8	8
.05	40	55	59	61	22
	50	68	70	54	40
.10	40	55	118	128	78
	50	68	126	134	103

Źródło: obliczenia własne.

**Tabela 4.** Empiryczna moc badanych testów dla asymetrycznego rozkładu trójkątnego w (%)

$\alpha$	$n$	Test			
		Davida–Hellwiga ( $\hat{\nu}_K$ )	Davida–Hellwiga ( $\hat{\nu}_{KR}$ )	Shapiro–Wilka ( $\hat{\nu}_w$ )	Kolmogorowa– Lillieforsa ( $\hat{\nu}_D$ )
.01	10	22	29	31	41
	20	20	42	75	35
	30	36	67	182	84
	40	36	66	390	131
	50	71	100	530	185
.05	10	22	110	124	105
	20	88	140	292	171
	30	107	203	443	228
	40	224	232	748	324
	50	282	287	820	437
.10	10	149	190	226	182
	20	88	237	451	300
	30	247	319	625	365
	40	224	350	861	490
	50	282	423	916	584

Źródło: obliczenia własne.



**Tabela 5.** Empiryczna moc badanych testów dla rozkładu dwumodalnego w (‰)

$\alpha$	$n$	Test			
		Davida–Hellwiga ( $\hat{\nu}_K$ )	Davida–Hellwiga ( $\hat{\nu}_{KR}$ )	Shapiro–Wilka ( $\hat{\nu}_w$ )	Kołmogorowa– Lillieforsa ( $\hat{\nu}_D$ )
.01	10	153	170	128	150
	20	341	432	699	295
	30	630	706	959	610
	40	726	784	999	856
	50	868	894	1000	946
.05	10	153	372	519	376
	20	619	691	950	729
	30	803	887	996	899
	40	947	948	1000	984
	50	981	981	1000	998
.10	10	469	514	721	527
	20	619	825	981	863
	30	926	943	1000	973
	40	947	968	1000	984
	50	981	991	1000	999

Źródło: obliczenia własne.

Uzyskane rezultaty przedstawione w tab. 2-5 pozwalają na sformułowanie następujących wniosków:

1. Moc testu Davida–Hellwiga dla rozważanych rozkładów alternatywnych symetrycznych jest zbliżona do mocy testu Kołmogorowa–Lillieforsa.

2. Moc testu Davida–Hellwiga jest również podobna do mocy testu Kołmogorowa przy asymetrycznym rozkładzie  $F_4$ , dla małych prób ( $n \leq 20$ ).

3. Zarówno test Davida–Hellwiga, jak i test Kołmogorowa–Lillieforsa mają znacznie mniejszą moc od testu Shapiro–Wilka przy wszystkich rozważanych rozkładach alternatywnych dla  $n > 20$ .

4. Dla małych prób  $n \leq 20$  test Davida–Hellwiga wykazuje pewną przewagę mocy nad testem Shapiro–Wilka przy symetrycznych rozkładach alternatywnych.

5. Dla rozkładu trójkątnego asymetrycznego moc wszystkich badanych testów jest niewielka.

## Literatura

- [1] Csorgo M., Guttman J., *On the empty cell test*, „Technometrics” 1962, vol. 4, s. 235-247.
- [2] David F.N., *Two combinatorial tests whether a sample has come from a given population*, „Biometrika” 1950, vol. 37, s. 97-110.
- [3] Domański C., *Testy statystyczne*, PWE, Warszawa 1990.
- [4] Domański C., Pruska K., *Nieklasyczne metody statystyczne*, PWE, Warszawa 2000.

- [5] Domański C., Tomaszewicz A., *Uwagi o teście zgodności Hellwiga*, „Przegląd Statystyczny” 1976, vol. 23, s. 453-459.
- [6] Domański C., Tomaszewicz A., *Moc testu Hellwiga przeciw niektórym hipotezom alternatywnym*, „Przegląd Statystyczny” 1980, vol. 27, s. 173-182.
- [7] Hellwig Z., *Test zgodności dla małej próby*, „Przegląd Statystyczny” 1965, vol. 12, s. 99-112.
- [8] Hellwig Z., *Uwagi o zorientowanych testach normalności*, *Wiadomości Statystyczne*” 1984, vol. 12, s. 16-17.
- [9] Kołczin V.F., Sewastianow B.A., Czistiakow V.P., *Sluczajnoje rozmieszczenija*, Nauka, Moskwa 1976.
- [10] Okamoto M., *On a non-parametric test*, „Osaka Mathematics Journal” 1952, vol. 3, s. 77-82.
- [11] Shapiro S.S., Wilk M.B., *An analysis of variance tests for normality (complete samples)*, „Biometrika” 1965, vol. 52, s. 591-611.

## THE PROPERTIES OF DAVID–HELLWIG TEST

**Summary:** The continuous growth of interest in non-parametric methods can be observed in the statistical literature. Generally, these methods are based on counting, rank and position statistics, as well as on a number and length of series. In this paper, also well-known test, based on a number of empty cells, is to be presented. The David–Hellwig consistency test will be discussed. Its empirical power will be shown, comparing to two classical tests: Kolmogorov–Lilliefors and Shapiro–Wilk’s for the hypothesis of the normality of a distribution.