# Interpretacja empiryczna efektywności Pitmana

Tadeusz Inglot

*Wydział Matematyki Politechniki Wrocławskiej*

**Streszczenie.** Badana jest interpretacja empiryczna efektywności Pitmana na przykładzie testowania jednostajności w dwuparametrowej rodzinie rozkładów beta. Dla ciągów mieszanek rozkładu hipotetycznego z ustaloną alternatywą efektywność Pitmana dobrze odpowiada empirycznym ilorazom liczebności prób.

**Klasyfikacja AMS:** 62G20, 62G10, 62F05
**Słowa kluczowe:** efektywność Pitmana, asymptotyczna efektywność względna, empiryczna efektywność względna, testowanie jednostajności, rozkład beta.

**1. Wstęp.** Efektywność Pitmana jest zwykle wyznaczana dla zagadnienia parametrycznego z parametrem rzeczywistym (np. Lehmann i Romano (2008)) albo dla podproblemu parametrycznego z parametrem rzeczywistym w problemie nieparametrycznym. Takie podejście nie pozwala porównywać wyników teoretycznych z empirycznymi dla wielowymiarowych zbiorów alternatyw.

Celem niniejszej notki jest zbadanie interpretacji empirycznej efektywności Pitmana dla problemu testowania jednostajności w dwuparametrowej rodzinie rozkładów beta. Najpierw podamy definicję efektywności Pitmana, którą przyjmujemy w tej notce.

Niech $\Gamma \subset \mathbb{R}^k,\ k \geqslant 1$, będzie zbiorem niepustym, $\mathcal{P} = \{P_\gamma : \gamma \in \Gamma\}$ rodziną rozkładów na przestrzeni mierzalnej $(\mathcal{X}, \mathcal{A})$ oraz $X_1, ..., X_n$ próbą z rozkładu $P \in \mathcal{P}$. Ustalmy $\gamma_0 \in \Gamma$. Testujemy hipotezę $H_0 : P = P_{\gamma_0}$ przeciwko $H_1 : P \neq P_{\gamma_0}$. Przypuśćmy, że chcemy porównać dwa testy prawostronne określone przez statystyki $T_n$, $V_n$. Dla $0 < \alpha < \beta < 1$ i alternatywy $P$ niech $N_T(\alpha, \beta, P)$ oznacza minimalną liczbę obserwacji taką, że dla wszystkich $n \geqslant N_T(\alpha, \beta, P)$ moc testu $T$ dla alternatywy $P$ na poziomie istotności $\alpha$ dla próby rozmiaru $n$ wynosi co najmniej $\beta$. Podobnie określamy $N_V(\alpha, \beta, P)$ dla testu $V$. Efektywność względna $T$ względem $V$ wyraża się wzorem

$$\mathcal{RE}_{TV}(\alpha, \beta, P) = \frac{N_V(\alpha, \beta, P)}{N_T(\alpha, \beta, P)}$$

Niech $\gamma(s) \in \Gamma,\ s \in [0, 1]$, będzie krzywą ciągłą w $\Gamma$, taką że $\gamma(0) = \gamma_0$. Załóżmy, że istnieje miara $\sigma$-skończona $\lambda$ na $(\mathcal{X}, \mathcal{A})$, taka że $P_{\gamma(s)} \ll \lambda$ dla $s \in [0, 1]$ oraz $H(P_{\gamma(s)}, P_{\gamma_0}) \xrightarrow{s \to 0^+} 0$, gdzie $H(P, Q)$ oznacza odległość Hellingera rozkładów $P$ i $Q$. Rodzinę $\{P_{\gamma(s)}\} = \{P_{\gamma(s)} : s \in [0, 1]\}$ nazywamy ścieżką.

**Definicja.** Niech $0 < \alpha < \beta < 1$ będą ustalone i niech $\{P_{\gamma(s)}\}$ będzie ścieżką. Jeśli istnieje granica

$$\lim_{s \to 0^+} \mathcal{RE}_{TV}(\alpha, \beta, P_{\gamma(s)}) = e^P_{TV}(\alpha, \beta, \{P_{\gamma(s)}\}) \in [0, \infty],$$

to nazywamy ją efektywnością Pitmana testu $T$ względem testu $V$ dla ścieżki $\{P_{\gamma(s)}\}$.

Poniżej podajemy wersję twierdzenia Pitmana w postaci wygodnej do zastosowania w rozdziale 2. Wprowadźmy następujące założenie o asymptotycznym zachowaniu się statystyki $W_n$ dla ścieżki $\{P_{\gamma(s)}\}$:

istnieją funkcje skalujące $\mu(s) \geqslant 0$, $\sigma(s) > 0$, $s \in [0,1]$, oraz dystrybuanta ciągła $G(x)$ takie, że

$$\lim_{n\to\infty} P^n_{\gamma_0}\left(\frac{W_n - \sqrt{n}\mu(0)}{\sigma(0)} \leqslant x\right) = G(x) \tag{1}$$

dla wszystkich $x \in \mathbb{R}$ oraz dla dowolnego ciągu $s_n \to 0$, $s_n > 0$, mamy

$$\lim_{n\to\infty} P^n_{\gamma(s_n)}\left(\frac{W_n - \sqrt{n}\mu(s_n)}{\sigma(s_n)} \leqslant x\right) = G(x) \tag{2}$$

dla wszystkich $x \in \mathbb{R}$.

Warunek (2) jest trochę słabszy niż jednostajna zbieżność względem $s$ (por. warunek (P1), Serfling (1980) lub warunek D, Noether (1955)). Rothe (1981) zaproponował trzy warunki zamiast (2). Jednym z nich jest ciągłość funkcji mocy względem $s$ w punkcie $s = 0$ dla każdego ustalonego $n$. Dowód poniższego twierdzenia jest modyfikacją znanych dowodów (por. Lehmann i Romano (2008), Nikitin (1995)). Dla kompletności podajemy go na końcu pracy. Zwykle założenia (1) i (2) są spełnione z $G(x) = \Phi(x)$ tj. dystrybuantą rozkładu standardowego normalnego. Jednak zarówno w sformułowaniu twierdzenia jak i jego dowodzie ten szczególny fakt jest nieistotny.

**Twierdzenie.** Załóżmy, że statystyki $T_n, V_n$ spełniają warunki (1) i (2) dla ścieżki $\{P_{\gamma(s)}\}$ z tą samą dystrybuantą $G(x)$, rosnącą na zbiorze $\{x : 0 < G(x) < 1\}$, funkcje $\sigma_T(s), \sigma_V(s)$ są ciągłe w $s = 0$, a $\mu_T(s), \mu_V(s)$ mają nieujemne pochodne w punkcie $s = 0$. Oznaczmy $c_T^P = (\mu_T'(0)/\sigma_T(0))^2$ oraz $c_V^P = (\mu_V'(0)/\sigma_V(0))^2$. Jeśli $\max\{c_T^P, c_V^P\} > 0$, to istnieje efektywność Pitmana $T$ względem $V$ dla ścieżki $\{P_{\gamma(s)}\}$, nie zależy od $\alpha$ oraz $\beta$ i wyraża się wzorem

$$e^P_{TV}(\{P_{\gamma(s)}\}) = e^P_{TV}(\alpha, \beta, \{P_{\gamma(s)}\}) = \left(\frac{\mu_T'(0)/\sigma_T(0)}{\mu_V'(0)/\sigma_V(0)}\right)^2 = \frac{c_T^P}{c_V^P}, \tag{3}$$

z umową $c/0 = \infty$.

**2. Przykład i interpretacja empiryczna.** W tym rozdziale zbadamy interpretację empiryczną efektywności Pitmana na przykładzie testowania jednostajności w rodzinie rozkładów beta. Niech

$$\mathcal{P} = \{P_\gamma : P_\gamma = P_{(p,q,\varepsilon)} = (1-\varepsilon)P_{1\,1} + \varepsilon P_{pq},\ \varepsilon \in [0,1],\ p \geqslant q > 0,\ \tau(p,q) \geqslant 0\},$$

gdzie $P_{pq}$ oznacza rozkład beta z parametrami $p, q$ oraz $\tau(p,q) = 2p^2 - 2pq - q^2 + 2p - q$. Zatem $\gamma_0 = (1,1,0)$ i $P_{\gamma_0} = P_{(1,1,0)} = P_{1\,1}$ jest rozkładem jednostajnym na odcinku [0,1]. Testujemy hipotezę prostą $H_0 : P = P_{1\,1}$. Rozważamy dwa testy (prawostronne) określone przez statystyki $V_n = \sqrt{n}(\overline{X} - 1/2)$, $T_n = (\sum_{i=1}^n (X_i^2 - 1/3))/\sqrt{n}$.

Przypomnijmy, że

$$E_{pq}X_1 = \frac{p}{p+q},\quad m_2 = E_{pq}X_1^2 = \frac{p(p+1)}{(p+q)(p+q+1)}, \tag{4}$$

$$m_4 = E_{pq}X_1^4 = \frac{p(p+1)(p+2)(p+3)}{(p+q)(p+q+1)(p+q+2)(p+q+3)}. \tag{5}$$

Zatem dla rozkładu $P_{pq}$ z $\tau(p,q) < 0$ mamy $E_{pq}X_1^2 = 1/3 + \tau(p,q)/3(p+q)(p+q+1) < 1/3$ i dla ścieżek leżących w obszarze $\tau(p,q) < 0$ statystyka $T_n$ nie spełnia (2), bo $\mu_T(s) < 0$. Dlatego rozważamy zbiór parametrów z ograniczeniem $\tau(p,q) \geqslant 0$.

Ustalmy rozkład $P_{pq} = P_{(p,q,1)} \in \mathcal{P}$, $(p,q) \neq (1,1)$, i niech $P_{\gamma(s)} = (1-s)P_{1\,1} + sP_{pq}$, $s \in [0,1]$. Mamy więc $\gamma(s) = (p,q,s)$ i ścieżka $\{P_{\gamma(s)}\}$ łączy odcinkiem liniowym (w przestrzeni rozkładów) $P_{1\,1}$ z $P_{pq}$. Nazywamy ją ścieżką liniową. Z twierdzenia Lapunowa i (4) wynika, że $V_n$ spełnia warunki (1) i (2) z $G(x) = \Phi(x)$, $\mu_V(s) = (1-s)/2 + sE_{pq}X_1 - 1/2 = s(p-q)/2(p+q)$ oraz $\sigma_V^2(s) = (1-s)/3 + sm_2 - (\mu_V(s)+1/2)^2$. Założenia twierdzenia 1 są dla tego testu spełnione i $\mu_V'(0) = (p-q)/2(p+q)$, $\sigma_V(0) = 1/\sqrt{12}$ oraz $c_V^P = 3(p-q)^2/(p+q)^2$. Podobnie z (4), (5) i twierdzenia Lapunowa wynika, że $T_n$ spełnia warunki (1) i (2) z $G(x) = \Phi(x)$, $\mu_T(s) = s\tau(p,q)/3(p+q)(p+q+1)$, $\sigma_T^2(s) = (1-s)/5 + sm_4 - (\mu_T(s)+1/3)^2$. Zatem założenia twierdzenia 1 dla tego testu są również spełnione i $\mu_T'(0) = \tau(p,q)/3(p+q)(p+q+1)$, $\sigma_T^2(0) = 4/45$ oraz $c_T^P = 5\tau^2(p,q)/4(p+q)^2(p+q+1)^2$. Z (3) wynika, że dla $p > q$ z $\tau(p,q) > 0$ efektywność Pitmana testu $T$ względem $V$ dla ścieżki liniowej wynosi

$$\mathcal{E}_{TV}^P = \mathcal{E}_{TV}^P(P_{pq}) = \frac{5(2p^2 - 2pq - q^2 + 2p - q)^2}{12(p-q)^2(p+q+1)^2} = \frac{5\tau^2(p,q)}{12(p-q)^2(p+q+1)^2}, \tag{6}$$

dla $p = q < 1$ efektywność jest równa $\infty$, a dla $p, q$ takich, że $\tau(p,q) = 0$ efektywność wynosi 0. Zauważmy, że dla $p, q$ leżących na prostej o równaniu $p - 2q + 1 = 0,\ p > 1$, mamy $\mathcal{E}_{TV}^P(P_{pq}) = 5/12$, a dla $p, q$ leżących na prostej $2p - q - 1 = 0,\ 1/2 < p < 1$, mamy $\mathcal{E}_{TV}^P(P_{pq}) = 5/3$. Dla ścieżek zawartych w tych prostych efektywność $e_{TV}^P$ jest taka sama i wynosi odpwiednio 5/12 i 5/3.

Porównajmy otrzymany wzór teoretyczny dla ścieżek liniowych z rzeczywistym zachowaniem się obu testów dla kilku alternatyw i poziomu istotności 0.05. Wyniki są podane w tabeli 1. Każdorazowo parametr $s$ został dobrany tak, aby moc była bliska 1/2 dla prób o umiarkowanych liczebnościach.

**Tabela 1.** Moce testów $V$ oraz $T$ (w %) oraz efektywności empiryczne dla wybranych alternatyw. $\alpha = 0.05$, 100 000 MC.

| alternatywa $p$ | $q$ | $s$ | $n$ | test $V$ | $T$ | efektyw. empiryczna | $\mathcal{E}_{TV}^P$ |
|---|---|---|---|---|---|---|---|
| 5 | 4 | 1 | 80 | 56 | 0 | | |
| | | | 30000 | 100 | 0 | $< 0.003$ | 0 |
| 4 | 3.15 | 0.9 | 100 | **63** | 03 | | |
| | | | 4000 | 100 | **63** | 0.025 | 0.026 |
| 6 | 4 | 0.5 | 100 | **54** | 15 | | |
| | | | 460 | 99 | **54** | 0.217 | 0.220 |
| 3 | 1 | 0.2 | 105 | **55** | 56 | | |
| | | | 100 | 54 | **55** | 1.05 | 1.067 |
| 0.6667 | 0.5 | 0.5 | 200 | **54** | 77 | | |
| | | | 100 | 36 | **54** | 2 | 2.074 |
| 0.55 | 0.5 | 0.9 | 640 | **58** | 100 | | |
| | | | 100 | 22 | **58** | 6.4 | 6.505 |
| 0.5 | 0.5 | 0.9 | 30000 | 09 | 100 | | |
| | | | 100 | 09 | 41 | $> 300$ | $\infty$ |

Dla pierwszej alternatywy moc testu $T$ wynosi 0.002 dla obu liczebności, co w tabeli zaznaczono jako 0. Widać niemal idealną zgodność efektywności empirycznych z ich teoretycznymi odpowiednikami dla wszystkich wartości $s$.

Rozważmy ścieżkę $\{P_{\gamma_1(s)}\}$ określoną przez krzywą $\gamma_1(s) = \gamma_1(s; p, q) = (1 - s + ps, 1 - s - qs + 2qs^2, 1)$ dla ustalonych $p \geqslant q > 0, \tau(p,q) > 0$ oraz $q < 3 + \sqrt{8}$. Oba parametry rozkładu beta ściągamy do 1, poruszając się po paraboli, bez mieszania z $P_{11}$. Warunek na $q$ gwarantuje, że krzywa jest zawarta w zbiorze parametrów (tj. $1 - s - qs + 2qs^2 > 0$ dla $s \in [0,1]$).

Z ciągłości gęstości rozkładu beta względem parametrów $p, q$ oraz z twierdzenia Lebesgue'a o zbieżności ograniczonej wynika, że odległość Hellingera $H(P_{\gamma_1(s)}, P_0)$ dąży do

0, gdy $s \to 0$ (parametry $p_1(s) = 1-s+ps$, $q_1(s) = 1-s-qs+2qs^2$ są odcięte od 0 jednostajnie względem $s$ co gwarantuje ograniczenie przez funkcję całkowalną). Z twierdzenia Lapunowa i (4) wynika, że

$$\mu_V(s) = \frac{1-s+ps}{2-2s+ps-qs+2qs^2} - \frac{1}{2} = \frac{(p+q)s-2qs^2}{2(2-2s+ps-qs+2qs^2)}, \quad s \in [0,1].$$

Zatem $\mu_V'(0) = (p+q)/4$. Podobnie jak poprzednio $\sigma_V^2(0) = 1/12$, więc $c_V^P = 3(p+q)^2/4$. Analogicznie, z twierdzenia Lapunowa i (4) mamy

$$\begin{aligned}\mu_T(s) &= \frac{(1-s+ps)(2-s+ps)}{(2-2s+ps-qs+2qs^2)(3-2s+ps-qs+2qs^2)} - \frac{1}{3} \\ &= \frac{(4p+5q+1)s}{3(2-2s+ps-qs+2qs^2)(3-2s+ps-qs+2qs^2)} + O(s^2), \quad s \to 0^+.\end{aligned}$$

Stąd $\mu_T'(0) = (4p+5q+1)/18$. Ponieważ $\sigma_T^2(0) = 4/45$, to $c_T^P = 5(4p+5q+1)^2/144$. Ostatecznie z (3)

$$e_{TV}^P(\{P_{\gamma_1(s)}\}) = \frac{5(4p+5q+1)^2}{108(p+q)^2}.$$

Dla przykładu weźmy $p = 6, q = 4$. Wtedy $e_{TV}^P(\{P_{\gamma_1(s;6,4)}\}) = 15/16 = 0.9375$. Wybierzmy kilka punktów na tej ścieżce dla $s = 1,\ 0.5,\ 0.2,\ 0.1,\ 0.05,\ 0.02$. W tabeli 2 podajemy efektywności empiryczne dla wybranych rozkładów na rozważanej ścieżce oraz wartości $\mathcal{E}_{TV}^P$, dla odpowiednich ścieżek liniowych, wyznaczone z (6). Ze względu na dużą czułość obu testów rozważamy mieszanki wybranych rozkładów z $P_{1\,1}$ postaci $(1-\varepsilon)P_{1\,1} + \varepsilon P_{p_1(s)\,q_1(s)}$ dla takich $\varepsilon$, aby liczebności prób nie były zbyt małe.

**Tabela 2.** Efektywności empiryczne testu $T$ względem $V$ dla wybranych rozkładów na ścieżce $\{P_{\gamma_1(s)}\}$. $\alpha = 0.05$, 100 000 MC.

| alternatywa | | | | | test | | efektyw. | |
|---|---|---|---|---|---|---|---|---|
| $s$ | $p_1(s)$ | $q_1(s)$ | $\varepsilon$ | $n$ | $V$ | $T$ | empiryczna | $\mathcal{E}_{TV}^P$ |
| 0.5 | 3.5 | 0.5 | 0.1 | 200 | **57** | 67 | | |
| | | | | 148 | 47 | **57** | 1.35 | 1.375 |
| 0.2 | 2 | 0.32 | 0.1 | 180 | **52** | 62 | | |
| | | | | 133 | 43 | **52** | 1.35 | 1.420 |
| 0.1 | 1.5 | 0.58 | 0.2 | 150 | **59** | 65 | | |
| | | | | 128 | 54 | **59** | 1.17 | 1.217 |
| 0.05 | 1.25 | 0.77 | 0.2 | 200 | **54** | 55 | | |
| | | | | 189 | 52 | **54** | 1.06 | 1.083 |
| 0.02 | 1.10 | 0.9032 | 0.2 | 200 | **52** | 50 | | |
| | | | | 207 | 53 | **52** | 0.97 | 0.996 |

Z tabeli 2 wynika, że efektywności empiryczne zgadzają się niemal idealnie z $\mathcal{E}_{TV}^P$ i dla rozkładów leżących na ścieżce relatywnie daleko od $P_{1\,1}$ efektywność $e_{TV}^P(\{P_{\gamma_1(s)}\}) = 0.9375$ nie odzwierciedla empirycznego zachowania się testów. Tylko dla alternatyw bardzo bliskich $P_{1\,1}$, $e_{TV}^P(\{P_{\gamma_1(s)}\})$ ma dobrą interpretację empiryczną. Ale wtedy obie efektywności różnią się niewiele. Jednak nawet w tym przypadku efektywności empiryczne są bliższe $\mathcal{E}_{TV}^P$ niż $e_{TV}^P$.

Elementarne obliczenia dają

$$\mathcal{E}_{TV}^P(P_{p_1(s)\,q_1(s)}) = \frac{5}{12}\left(\frac{4p+5q+1+o(s)}{3(p+q)+o(s)}\right)^2.$$

Stąd $\lim\limits_{s\to 0^+} \mathcal{E}_{TV}^P(P_{p_1(s)\,q_1(s)}) = e_{TV}^P(\{P_{\gamma_1(s)}\})$ co widać w ostatniej kolumnie tabeli 2.

Ponieważ w powyższym przykładzie konieczne było mieszanie rozkładów z $P_{1\,1}$, rozważymy dwie inne ścieżki dane przez krzywe $\gamma_2(s) = (1+2s+s^2, 1+s+s^2, 1)$ oraz $\gamma_3(s) = (1-s/2+s^2/2, 1-2s/3, 1)$. Łaczą one odpowiednio $P_{1\,1}$ z $P_{43}$ i $P_{1\,1}$ z $P_{1\,1/3}$ bez mieszania z $P_{1\,1}$. Analogiczne obliczenia z użyciem wzorów (4) i (5) dają dla $\gamma_2(s)$: $\mu'_V(0) = 1/4$, $\mu'_T(0) = 1/6$ i w rezultacie $c_V^P = 3/4$, $c_T^P = 5/16$, $e_{TV}^P(\{P_{\gamma_2(s)}\}) = 5/12 \approx 0.4167$. Natomiast dla $\gamma_3(s)$ odpowiednio: $\mu'_V(0) = 1/24$, $\mu'_T(0) = 2/27$ oraz $c_V^P = 1/48$, $c_T^P = 5/81$, $e_{TV}^P(\{P_{\gamma_3(s)}\}) = 80/27 \approx 2.963$.

Wybierzmy po 5 punktów na tych ścieżkach, kładąc $s = 1, 0.5, 0.2, 0.1, 0.05$. W tabeli 3 podajemy efektywności empiryczne dla drugiej ścieżki. W tym przypadku nie jest potrzebne mieszanie z $P_{1\,1}$, aby moc znalazła się w środkowym zakresie dla prób o umiarkowanej liczebności. Zaznaczamy to w tabeli w kolumnie $\varepsilon$.

**Tabela 3.** Efektywności empiryczne testu $T$ względem $V$ dla wybranych rozkładów na ścieżce $\{P_{\gamma_2(s)}\}$. $\alpha = 0.05$, 100 000 MC.

| alternatywa | | | | | test | | efektyw. | |
|---|---|---|---|---|---|---|---|---|
| $s$ | $p_2(s)$ | $q_2(s)$ | $\varepsilon$ | $n$ | $V$ | $T$ | empiryczna | $\mathcal{E}_{TV}^P$ |
| 1 | 4 | 3 | 1 | 50 | **57** | 05 | | |
| | | | | 497 | 100 | **57** | 0.101 | 0.104 |
| 0.5 | 2.25 | 1.75 | 1 | 70 | **59** | 18 | | |
| | | | | 290 | 100 | **59** | 0.241 | 0.250 |
| 0.2 | 1.44 | 1.24 | 1 | 200 | **58** | 27 | | |
| | | | | 582 | 95 | **58** | 0.344 | 0.354 |
| 0.1 | 1.21 | 1.11 | 1 | 600 | **58** | 29 | | |
| | | | | 1600 | 92 | **58** | 0.375 | 0.387 |
| 0.05 | 1.1025 | 1.0525 | 1 | 2000 | **56** | 30 | | |
| | | | | 5015 | 89 | **56** | 0.399 | 0.402 |

W kolejnej tabeli podajemy efektywności empiryczne dla trzeciej ścieżki. Dla dwóch ostatnich wartości wzięto 10000 powtórzeń MC, a w pozostałych przypadkach 100 000 powtórzeń.

**Tabela 4.** Efektywności empiryczne testu $T$ względem $V$ dla wybranych rozkładów na ścieżce $\{P_{\gamma_3(s)}\}$. $\alpha = 0.05$.

| alternatywa | | | | | test | | efektyw. | |
|---|---|---|---|---|---|---|---|---|
| $s$ | $p_3(s)$ | $q_3(s)$ | $\varepsilon$ | $n$ | $V$ | $T$ | empiryczna | $\mathcal{E}_{TV}^P$ |
| 1 | 1 | 0.3333 | 0.2 | 100 | **53** | 64 | | |
| | | | | 73 | 44 | **54** | 1.37 | 1.437 |
| 0.5 | 0.875 | 0.6667 | 1 | 70 | **62** | 74 | | |
| | | | | 49 | 50 | **62** | 1.43 | 1.496 |
| 0.2 | 0.92 | 0.8667 | 1 | 1000 | **50** | 72 | | |
| | | | | 535 | 33 | **50** | 1.87 | 1.936 |
| 0.1 | 0.955 | 0.9333 | 1 | 7000 | **51** | 79 | | |
| | | | | 3120 | 30 | **51** | 2.24 | 2.295 |
| 0.05 | 0.97625 | 0.9667 | 1 | 35000 | **48** | 81 | | |
| | | | | 13650 | 26 | **48** | 2.56 | 2.58 |

Wyniki przedstawione w tabelach 3 i 4 potwierdzają spostrzeżenia poczynione dla ścieżki $\{P_{\gamma_1(s)}\}$. Widać, że efektywność Pitmana istotnie zależy od ścieżki i wartości znacznie różnią się między sobą (w naszych przykładach: 0.9375, 0.4167, 2.963). Co więcej, ich interpretacja empiryczna jest poprawna tylko dla alternatyw leżących na danej ścieżce i odpowiadających bardzo małym $s$. Natomiast efektywność $\mathcal{E}_{TV}^P$ dla ścieżek liniowych ma dobrą interpretację empiryczną w każdym punkcie ścieżki.

Rozważmy jeszcze jedną ścieżkę określoną przez krzywą $\gamma_4(s) = (1+s, 1+0.5s-1.18s^2, 1)$. Krzywe $\gamma_3(s)$ i $\gamma_4(s)$ są styczne w $s=0$ do prostej o równaniu $p-2q+1=0$. Podobne obliczenia dają $e^P_{TV}(\{P_{\gamma_4(s)}\}) = 5/12$. Ścieżka $\{P_{\gamma_1(s;6,4)}\}$ przecina się z $\{P_{\gamma_4(s)}\}$ dla rozkładu $P_{2\,0.32}$. Można mu więc przypisać liczby $5/12$ oraz $15/16$ jako efektywności (teoretyczne). Jednak obie liczby niewiele mają wspólnego z efektywnością empiryczną wynoszącą 1.35 (por. tabela 2). W tabeli 5 przedstawiamy efektywności empiryczne dla $\{P_{\gamma_4(s)}\}$ dla kilku wartości $s$. Także w tym przykładzie efektywności empiryczne są lepiej przybliżane przez $\mathcal{E}^P_{TV}$ niż przez $e^P_{TV}$ wynoszącą $5/12$.

Ścieżka $\{P_{\gamma_4(s)}\}$ przecina się ze ścieżką $\{P_{\gamma_5(s)}\}$ wyznaczoną przez odcinek liniowy $\gamma_5(s) = (1+s, 1, 1)$ dla rozkładu $P_{84/59\,1}$. Mamy $e^P_{TV}(\{P_{\gamma_5(s)}\}) = 20/27 \approx 0.741$ oraz $e^P_{TV}(\{P_{\gamma_4(s)}\}) = 5/12$ i obie te liczby możnaby przypisać temu rozkładowi jako efektywność. Natomiast, $\mathcal{E}^P_{TV}(P_{84/59\,1}) \approx 0.835$, a efektywność empiryczna wynosi 0.800 i jest gorzej przybliżana przez efektywności teoretyczne dla obu ścieżek niż przez $\mathcal{E}^P_{TV}$.

**Tabela 5.** Efektywności empiryczne testu $T$ względem $V$ dla wybranych rozkładów na ścieżce $\{P_{\gamma_4(s)}\}$. $\alpha = 0.05$, 100 000 MC.

| alternatywa | | | | | test | | efektyw. | |
|---|---|---|---|---|---|---|---|---|
| $s$ | $p_4(s)$ | $q_4(s)$ | $\varepsilon$ | $n$ | $V$ | $T$ | empiryczna | $\mathcal{E}^P_{TV}$ |
| 0.5 | 1.5 | 0.955 | 0.5 | 100 | **61** | 56 | | |
| | | | | 116 | 67 | **61** | 0.862 | 0.900 |
| 0.2 | 1.2 | 1.0528 | 1 | 250 | **56** | 40 | | |
| | | | | 405 | 74 | **56** | 0.617 | 0.637 |
| 0.1 | 1.1 | 1.0382 | 1 | 1300 | **56** | 36 | | |
| | | | | 2480 | 81 | **56** | 0.524 | 0.536 |

Prawdziwy (nieznany) rozkład $P \neq P_{1\,1}$ (przy prawdziwości $H_1$) leży na wielu ścieżkach. Nie można mu zatem przypisać jednej liczby interpretowanej jako efektywność i rozumianej jako iloraz liczebności prób potrzebnych do osiągnięcia tej samej mocy. Nawet dla alternatyw bliskich $P_{1\,1}$ (w sensie odległości Hellingera) efektywność Pitmana może znacznie odbiegać od efektywności empirycznej, gdyż zależy to od przebiegu krzywej (ścieżki) i wielkości parametru $s$ odpowiadającego temu rozkładowi na danej ścieżce. Jest tak, gdy krzywa przebiega blisko $P_{1\,1}$ w sensie odległości Hellingera dla dużych wartości parametru $s$. Na przykład, dla krzywej $\gamma_6(s) = (1-s+11s^2/10, 1-2s+2s^2, 1)$ mamy $e^P_{TV}(\{P_{\gamma_6(s)}\}) = 5/3$, ale dla $s=1$ mamy $P_{\gamma_6(1)} = P_{1.1\,1}$ oraz $H(P_{1.1\,1}, P_{1\,1}) = 0.048$ i empiryczną efektywność 0.735 bliską $\mathcal{E}^P_{TV}(P_{1.1\,1}) \approx 0.765$.

Weźmy jeszcze jedną ścieżkę wyznaczoną przez $\gamma_7(s) = (2+s, 1+s, s)$. Dla dowolnego

**Tabela 6.** Efektywności empiryczne $T$ względem $V$ dla wybranych rozkładów na ścieżce $\{P_{\gamma_7(s)}\}$. $\alpha = 0.05$, 100 000 MC.

| alternatywa | | | | | test | | efektyw. | |
|---|---|---|---|---|---|---|---|---|
| $s$ | $p_7(s)$ | $q_7(s)$ | $\varepsilon$ | $n$ | $V$ | $T$ | empiryczna | $\mathcal{E}^P_{TV}$ |
| 1 | 3 | 2 | 1 | 30 | 64 | 29 | | |
| | | | | 76 | 98 | 64 | 0.39 | 0.417 |
| 0.5 | 2.5 | 1.5 | 0.5 | 80 | 63 | 45 | | |
| | | | | 127 | 81 | 63 | 0.63 | 0.651 |
| 0.2 | 2.2 | 1.2 | 0.2 | 300 | 55 | 47 | | |
| | | | | 380 | 63 | 55 | 0.79 | 0.817 |
| 0.1 | 2.1 | 1.1 | 0.1 | 1000 | 54 | 48 | | |
| | | | | 1205 | 60 | 54 | 0.83 | 0.876 |
| 0.05 | 2.05 | 1.05 | 0.05 | 4000 | 55 | 50 | | |
| | | | | 4440 | 59 | 55 | 0.90 | 0.907 |

$s_n \to 0$ mamy ciąg alternatyw postaci $(1-s_n)P_{1\,1} + s_n P_{2+s_n\,1+s_n}$. Każdemu $s$ odpowiada teraz inna alternatywa w mieszance. Ciąg gęstości rozkładów $P_{2+s_n\,1+s_n}$ jest jednostajnie ograniczony i jednostajnie odcięty od 0 i jest zbieżny do gęstości rozkładu $P_{2\,1}$. Efektywność Pitmana dla tej ścieżki wynosi $e^P_{TV}(\{P_{\gamma_7(s)}\}) = 15/16$ i jest równa $\mathcal{E}^P_{TV}(P_{2\,1})$ dla rozkładu granicznego $P_{2\,1}$. Jednak znowu efektywności empiryczne są dla każdego $n$ bliższe efektywnościom $\mathcal{E}^P_{TV}(P_{2+s_n\,1+s_n})$ niż $e^P_{TV}(\{P_{\gamma_7(s)}\})$ co widać w tabeli 6. Łatwo sprawdzić, że $\mathcal{E}^P_{TV}(P_{2+s_n\,1+s_n}) \to 15/16 = e^P_{TV}(\{P_{\gamma_7(s)}\})$ gdy $n \to \infty$, a z tabeli 6 widać, że efektywności empiryczne również dążą do 15/16 gdy $s \to 0$.

**Wniosek.** Interpretacja empiryczna efektywności Pitmana dla dowolnych ścieżek napotyka na trudności. Natomiast, dla ścieżek liniowych efektywność Pitmana dobrze przybliża empiryczną efektywność względną.

### Literatura


Lehmann, E. L. i Romano, J. P. (2008), *Testing Statistical Hypotheses*, Springer Texts in Statistics, Springer, New York, wyd. 3.

Nikitin, Y. (1995), *Asymptotic Efficiency of Nonparametric Tests*, Cambridge University Press, Cambridge.

Noether, G. E. (1955), On a theorem of Pitman, *Ann. Math. Statist.*, **26**, 64-68.

Rothe, G. (1981), Some properties of the asymptotic relative Pitman efficiency, *Ann. Statist.*, **9**, 663-669.

Serfling, R., J. (1980), *Approximation Theorems of Mathematical Statistics*, Wiley, New York.


**Dowód twierdzenia.** Ustalmy $0 < \alpha < \beta < 1$ i ciąg $s_n \to 0^+$ i oznaczmy przez $t_{\alpha n}, v_{\alpha n}$ dokładne wartości krytyczne obu porównywanych testów.

(i) Wykażemy, że $N_T(\alpha, \beta, P_{\gamma(s_n)}) \to \infty$, gdy $n \to \infty$.
Niech $p_0(x), p_s(x)$ oznaczają gęstości $P_{\gamma_0}, P_{\gamma(s)}$ względem $\lambda$. Oznaczmy $\kappa_{nT} = N_T(\alpha, \beta, P_{\gamma(s_n)})$. Dla czytelności będziemy opuszczać indeks $T$ i pisać krócej $\kappa_n$. Przez $p_{n0}, p_{ns_n}$ oznaczamy gęstość łączną próby z rozkładu odpowiednio $P_{\gamma_0}, P_{\gamma(s_n)}$. Dla $A_n = \{T_{\kappa_n} > t_{\alpha\kappa_n}\}$ mamy z nierówności Schwarza

$$\begin{aligned} 0 < \beta - \alpha &\leqslant \int_{A_n} (p_{\kappa_n s_n} - p_{\kappa_n 0}) d\lambda^{\kappa_n} \\ &= \int_{A_n} (\sqrt{p_{\kappa_n s_n}} - \sqrt{p_{\kappa_n 0}})^2 d\lambda^{\kappa_n} + 2\int_{A_n} (\sqrt{p_{\kappa_n s_n}} - \sqrt{p_{\kappa_n 0}})\sqrt{p_{\kappa_n 0}} d\lambda^{\kappa_n} \\ &\leqslant H^2(P^{\kappa_n}_{\gamma(s_n)}, P^{\kappa_n}_{\gamma_0}) + 2\sqrt{\alpha} H(P^{\kappa_n}_{\gamma(s_n)}, P^{\kappa_n}_{\gamma_0}). \end{aligned} \tag{7}$$

Gdyby $\kappa_n$ zawierał podciąg ograniczony, to, z postaci odległości Hellingera dla produktów rozkładów tj. $H^2(P^{\kappa_n}_{\gamma(s_n)}, P^{\kappa_n}_{\gamma_0}) = 2 - 2(1 - H^2(P_{\gamma(s_n)}, P_{\gamma_0})/2)^{\kappa_n}$, prawa strona (7) byłaby zbieżna do zera dla tego podciągu, co przeczyłoby wyborowi $\alpha$ i $\beta$. Zatem $\kappa_n \to \infty$.

(ii) Dla testu $T$ mamy $P^{\kappa_n}_{\gamma_0}(T_{\kappa_n} > t_{\alpha\kappa_n}) \leqslant \alpha \leqslant P^{\kappa_n}_{\gamma_0}(T_{\kappa_n} \geqslant t_{\alpha\kappa_n})$. Ponieważ $G(x)$ jest rosnąca i ciągła oraz $0 < \alpha < 1$, to z (1) wynika

$$\frac{t_{\alpha\kappa_n} - \sqrt{\kappa_n}\mu_T(0)}{\sigma_T(0)} \to G^{-1}(1-\alpha) = z_\alpha.$$

Zatem $t_{\alpha\kappa_n} = \sqrt{\kappa_n}\mu_T(0) + \sigma_T(0) z_\alpha + o(1)$.

(iii) Oznaczmy $z_\beta = G^{-1}(1-\beta)$. Pokażemy, że $s_n\sqrt{\kappa_n}$ jest zbieżny do $(z_\alpha - z_\beta)/\sqrt{c^P_T}$, gdy $c^P_T > 0$, oraz do $\infty$, gdy $c^P_T = 0$.

Dla $c^P_T > 0$, przypuśćmy przeciwnie, że dla pewnego rosnącego ciągu liczb naturalnych $k_n$ jest $s_{k_n}\sqrt{\kappa_{k_n}} \to C < (z_\alpha - z_\beta)/\sqrt{c^P_T}$ oraz $\kappa_{k_n}$ jest rosnący. Z definicji $\kappa_n$

$$\beta \leqslant P^{\kappa_n}_{\gamma(s_n)}(T_{\kappa_n} > t_{\alpha\kappa_n}) = P^{\kappa_n}_{\gamma(s_n)}(T_{\kappa_n} > \sqrt{\kappa_n}\mu_T(0) + \sigma_T(0)z_\alpha + o(1))$$
$$= P^{\kappa_n}_{\gamma(s_n)}\left(\frac{T_{\kappa_n} - \sqrt{\kappa_n}\mu_T(s_n)}{\sigma_T(s_n)} > \frac{(\mu_T(0) - \mu_T(s_n))\sqrt{\kappa_n} + \sigma_T(0)z_\alpha + o(1)}{\sigma_T(s_n)}\right).$$

Stąd dla podciągu $k_n$

$$\beta \leqslant P^{\kappa_{k_n}}_{\gamma(s_{k_n})}\left(\frac{T_{\kappa_{k_n}} - \sqrt{\kappa_{k_n}}\mu_T(s_{k_n})}{\sigma_T(s_{k_n})} > -\frac{\mu_T(0) - \mu_T(s_{k_n})}{s_{k_n}}\frac{s_{k_n}\sqrt{\kappa_{k_n}}}{\sigma_T(s_{k_n})} + \frac{\sigma_T(0)z_\alpha}{\sigma_T(s_{k_n})} + o(1)\right). \quad (8)$$

Określmy ciąg $\vartheta_n$ następująco: $\vartheta_{\kappa_{k_n}} = s_{k_n}$ oraz $\vartheta_j = s_j$ dla $j \neq \kappa_{k_n}$. Wtedy $\vartheta_n \to 0$ i z (2) zastosowanego do $\vartheta_n$ dla podciągu $\kappa_{k_n}$ oraz z istnienia $\mu'_T(0)$, po przejściu do granicy w (8), dostajemy $\beta \leqslant 1 - G(-C\sqrt{c^P_T} + z_\alpha)$. To daje sprzeczność, gdyż $z_\alpha - C\sqrt{c^P_T} > z_\beta$.

Analogicznie,przypuśćmy, że dla pewnego rosnącego ciągu liczb naturalnych $k_n$ jest $s_{k_n}\sqrt{\kappa_{k_n}} \to C > (z_\alpha - z_\beta)/\sqrt{c^P_T}$. Z minimalności $\kappa_n$ mamy

$$\beta > P^{\kappa_n - 1}_{\gamma(s_n)}(T_{\kappa_n - 1} > t_{\alpha\kappa_n - 1})$$
$$= P^{\kappa_n - 1}_{\gamma(s_n)}\left(\frac{T_{\kappa_n - 1} - \sqrt{\kappa_n - 1}\mu_T(s_n)}{\sigma_T(s_n)} > -\frac{(\mu_T(0) - \mu_T(s_n))\sqrt{\kappa_n - 1} + \sigma_T(0)z_\alpha + o(1)}{\sigma_T(s_n)}\right).$$

Dla podciągu $k_n$

$$\beta \geqslant P^{\kappa_{k_n} - 1}_{\gamma(s_{k_n})}\left(\frac{T_{\kappa_{k_n} - 1} - \sqrt{\kappa_{k_n} - 1}\mu_T(s_{k_n})}{\sigma_T(s_{k_n})} > -\frac{\mu_T(0) - \mu_T(s_{k_n})}{s_{k_n}}\frac{s_{k_n}\sqrt{\kappa_{k_n} - 1}}{\sigma_T(s_{k_n})} + z_\alpha + o(1)\right). (9)$$

Określmy ciąg $\vartheta_n$ nastepująco: $\vartheta_{\kappa_{k_n} - 1} = s_{k_n}$ oraz $\vartheta_j = s_j$ dla $j \neq \kappa_{k_n} - 1$. Wtedy $\vartheta_n \to 0$ i z (2) zastosowanego dla podciągu $\kappa_{k_n} - 1$ po przejściu do granicy w (9), analogicznie jak powyżej, dostajemy $\beta \geqslant 1 - G(-C\sqrt{c^P_T} + z_\alpha)$, co daje sprzeczność z nierównością $z_\alpha - C\sqrt{c^P_T} < z_\beta$.

Jeśli $c^P_T = 0$, to przypuśćmy, że dla pewnego rosnącego ciągu liczb naturalnych $k_n$ jest $s_n\sqrt{\kappa_n} \to C < \infty$. Powtarzając analogiczne rozumowanie jak w (8), dostajemy nierówność $\beta \leqslant 1 - G(z_\alpha)$, co również daje sprzeczność.

(iv) Tak samo jak w (iii) dowodzimy, że $s_n\sqrt{\kappa_{nV}} \to (z_\alpha - z_\beta)/\sqrt{c^P_V}$, gdy $c^P_V > 0$, lub $\infty$, gdy $c^P_V = 0$. Jeśli $c^P_T, c^P_V > 0$, to stąd i z (iii) natychmiast otrzymujemy

$$\frac{N_V(\alpha, \beta, P_{\gamma(s_n)})}{N_T(\alpha, \beta, P_{\gamma(s_n)})} = \frac{s_n^2\kappa_{nV}}{s_n^2\kappa_{nT}} \longrightarrow \frac{c_T}{c_V} = \left(\frac{\mu'_T(0)/\sigma_T(0)}{\mu'_V(0)/\sigma_V(0)}\right)^2.$$

Jeśli $c_T = 0, c_V > 0$, to $s_n\sqrt{\kappa_{nT}} \to \infty$, a $s_n\sqrt{\kappa_{nV}} \to (z_\alpha - z_\beta)/c_V$. Zatem $\kappa_{nV}/\kappa_{nT} = s_n^2\kappa_{nV}/s_n^2\kappa_{nT} \to 0$. Podobnie dla $c_T > 0, c_V = 0$ dostajemy $\kappa_{nV}/\kappa_{nT} \to \infty$. To kończy dowód twierdzenia. □

**Uwaga 2.** W dowodzie korzystaliśmy z własności, że dla liczebności próby $\kappa_{nT} = N_T(\alpha, \beta, P_{\gamma(s)})$ moc testu $T$ wynosi co najmniej $\beta$ oraz dla liczebności $\kappa_{nT} - 1$ moc $T$ jest mniejsza od $\beta$. Wszystkie liczebności $\tilde{N}_n$ mające tę samą własność mają taką samą asymptotykę jak $\kappa_{nT}$ (por. pojęcie ciągu $\beta$-PEF w pracy Rothe (1981), str. 663).

**Uwaga 3.** Rothe (1981) pokazał, że zbieżność mocy testu $W$ do $\beta > \alpha$ dla ciągów $s_n \to 0$ określonej postaci uzupełniona dwoma warunkami pociąga zbieżność mocy do $\beta$ dla wszystkich ciągów $s_n \to 0$ dla próby o liczebności $N_W(\alpha, \beta, P_{\gamma(s_n)})$ (czyli daje krok (iii) w powyższym dowodzie). Jednym z tych warunków jest ciągłość funkcji mocy względem $s$ w punkcie $s = 0$ dla każdej ustalonej liczebności próby $n$ (condition B, str. 665, Rothe (1981)). Dla przykładu rozważanego przez nas w rozdziale 2 ten warunek daje się dość łatwo sprawdzić, ale w ogólnym przypadku niekoniecznie. Manipulowanie podciągami w dowodzie kroku (iii) jest podobne do postępowania w dowodzie lematu Kallenberga, który podaje warunki wystarczające istnienia efektywności Kallenberga (efektywności pośredniej).