Journal of Polish CIMEEAC
Faculty of Mechanical Engineering and Ship Technology
Institute of Ocean Engineering and Ship Technology
Department of Marine Power Plants

ZNACZENIE MOCY TESTU STATYSTYCZNEGO W ANALIZIE DIAGNOSTYCZNEJ OKRĘTOWEGO SILNIKA O ZS

Patrycja Puzdrowska

Politechnika Gdańska
Wydział Inżynierii Mechanicznej i Okrętownictwa
Zakład Siłowni Okrętowych
80-233 Gdańsk ul. G. Narutowicza 11/12
e-mail: patpuzdr@pg.edu.pl

ABSTRAKT

W artykule przedstawiono sposób interpretacji błędów pierwszego (α) i drugiego (β) rodzaju oraz wartości krytycznej F_{kr} rozkładu Fishera-Snedecora. Omówiono dwa sposoby weryfikacji istotności wpływu parametru wejściowego na wyjściowy. Przeanalizowano znaczenie mocy testu statystycznego M w diagnostyce parametrycznej okrętowych silników o ZS. Przeprowadzono analizę wpływu wartości poziomu istotności na wartość krytyczną statystyki i odwrotnie, z użyciem kalkulatorów statystycznych. Oceniono czy zwiększanie mocy testu statystycznego i poziomu istotności ma uzasadnienie podczas wnioskowania statystycznego.

Słowa kluczowe: moc testu, statystyka F rozkładu Fishera-Snedecora, kalkulator statystyczny, diagnostyka parametryczna

1. Wstęp

Aby opracować metodykę diagnozowania elementów ograniczających komorę spalania okrętowego silnika o ZS wraz z aparaturą paliwową oraz kanałami powietrza dolotowego i spalin wylotowych na podstawie szybkozmiennej temperatury spalin wylotowych istotne są przede wszystkim dwa elementy. Po pierwsze należy opracować odpowiednią technologię jej pomiaru, uwzględniającą szeroko rozumiane warunki realizacji badań oraz niepewność pomiarową. W przypadku dynamicznych pomiarów temperatury spalin istotne jest zastosowanie termopar o jak najmniejszej inercyjności [Korczewski i Puzdrowska, 2015]. Jej przebieg czasowy, uzyskany podczas badań na obiekcie rzeczywistym, jakim jest okrętowy silnik o ZS, jest obarczony licznymi zakłóceniami mierzalnymi i niemierzalnymi, które wynikają z wpływów otoczenia oraz sieci pomiarowej. Należy w pierwszej kolejności poddać go odpowiedniej obróbce matematycznej pozwalającej odtworzyć przebieg najbardziej zbliżony do rzeczywistego [Puzdrowska, 2018]. Aby jednak tak uzyskany sygnał był diagnoście przydatny należy go poddać analizie statystycznej i merytorycznej wykorzystując do tego celu odpowiednie narzędzia [Piotrowski, 1976; Pabis, 1985; Korzyński, 2017]. W przypadku statystyki F rozkładu Fishera-Snedecora[1] należy sformułować odpowiednią hipotezę zerową oraz wybrać sposób weryfikacji istotności wpływu czynnika wejściowego na wyjściowy. Oba zaproponowane sposoby wiążą się nierozerwalnie z wartością krytyczną statystyki F_{kr} oraz z poziomem istotności α. Z kolei obie te wartości wpływają na moc testu

[1] W dalszej części artykułu nazywana krócej "statystyką F"

statystycznego M, rozumianego jako miara skuteczności testu do wykrywania fałszywej hipotezy zerowej. Jednakże należy pamiętać, że zwiększająca się moc testu statystycznego powoduje jednoczesny spadek możliwości popełnienia błędu drugiego rodzaju β, czyli przyjęcia hipotezy zerowej, gdy jest ona fałszywa, ale także wzrost możliwości popełnienia błędu pierwszego rodzaju wiążącego się z przyjętym arbitralnie poziomem istotności, tzn. prawdopodobieństwem odrzucenia hipotezy zerowej wtedy, gdy jest ona prawdziwa - tab. 1.

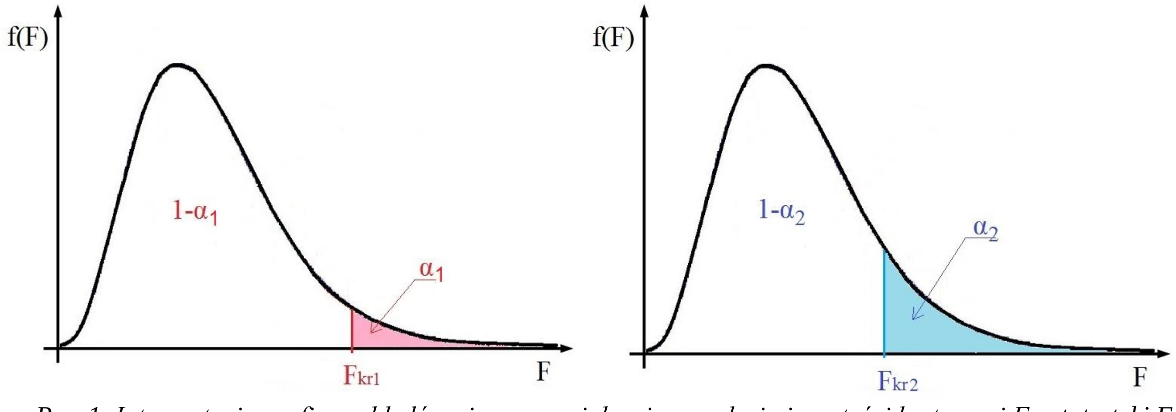
Tab. 1. Błędy popełniane przy ocenie hipotez statystycznych [Korzyński, 2017]

Hipoteza	Przyjęta	Odrzucona
Prawdziwa	Prawidłowa decyzja (1- α)	Błąd pierwszego rodzaju (α)
Fałszywa	Błąd drugiego rodzaju (β)	Poprawna decyzja (1-β)

2. Moc testu statystycznego

Ze względu na charakter prowadzonych badań, autorka niniejszego artykułu stosowała w nich program statyczny randomizowany kompletny, który pozwolił na ocenę istotności wpływu jednego czynnika wejściowego na czynnik wyjściowy [Korzyński, 2017; Puzdrowska 2019 i 2021a]. Hipoteza zerowa, formułowana z góry i weryfikowana w badaniach statystycznych, zakłada brak wpływu czynnika wejściowego na czynnik wynikowy. Wpływ czynnika wejściowego uważa się za istotny, gdy wartość obliczeniowa przyjętej statystyki jest równa lub większa od wartości krytycznej, podawanej w tablicach dla przyjętej wartości poziomu istotności i liczby stopni swobody. Uznano, że w prowadzonych badaniach diagnostycznych najlepiej jest przyjąć statystykę F rozkładu Fishera - Snedecora, ponieważ spełnione zostały warunki zastosowania jednego z jednostronnych testów parametrycznych [Polański, 1984]. W prowadzonych badaniach założono z góry, że wyniki pomiarów wszystkich parametrów kontrolnych można zamodelować jako zmienne losowe o rozkładzie normalnym, o określonej wariancji będącej miarą rozrzutu ich wartości wokół wartości średniej. Przyjęto także, że wariancje zmiennych losowych są równe lub zbliżone co do wartości, a zastosowane testy parametryczne dotyczące wariancji charakteryzują się jednostronnym obszarem krytycznym.

Jak widać na rysunku 1, gdy przyjęty zostanie poziom istotności większy ($\alpha_1<\alpha_2$), wówczas należy się liczyć z tym, że obszar 1-α zmniejszy się . Zatem zwiększenie wartości α (większe prawdopodobieństwo odrzucenia hipotezy zerowej, gdy jest ona prawdziwa) wiąże się z jednoczesnym zmniejszeniem możliwości popełnienia błędu II rodzaju wynoszącego β=1-α (mniejsze prawdopodobieństwo przyjęcia hipotezy zerowej, gdy jest fałszywa) i odwrotnie. Wartość krytyczna statystyki F_{kr} także rośnie wraz ze spadkiem przyjętego α ($F_{kr1}>F_{kr2}$).



Rys. 1. Interpretacja graficzna błędów pierwszego i drugiego rodzaju i wartości krytycznej F_{kr} statystyki F rozkładu Fishera - Snedecora, gdzie: α - prawdopodobieństwo popełnienia błędu I rodzaju, f(F) - funkcja gęstości prawdopodobieństwa rozkładu Fishera – Snedecora

Po wyznaczeniu wartości krytycznej statystyki F_{kr} możliwe jest dalsze postępowanie zgodnie z jednym z dwóch sposobów weryfikacji istotności wpływu czynnika wejściowego na wynikowy [Polański, 1984]:

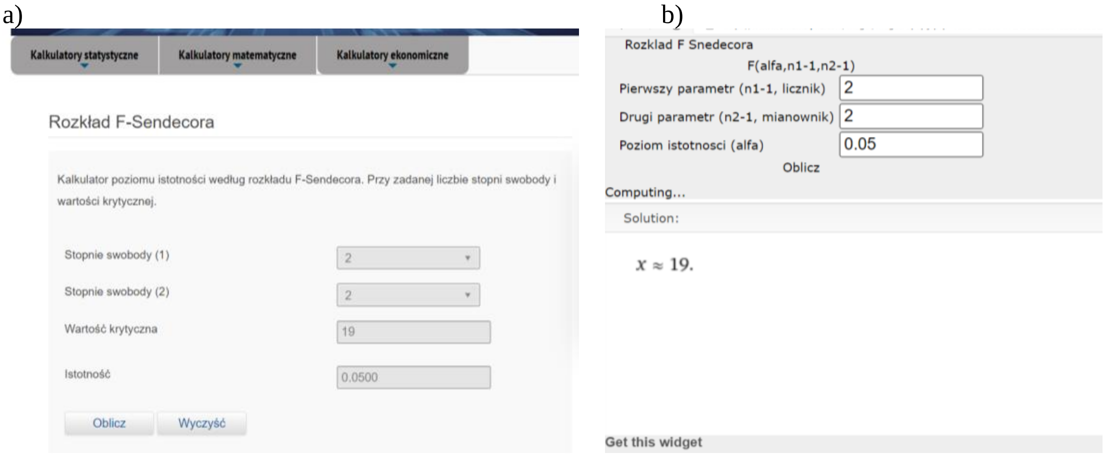
- weryfikacja dzięki porównaniu wartości statystyki: obliczeniowej F_{obl} i krytycznej F_{kr},
- weryfikacja na drodze porównania poziomu istotności $\alpha(F_{obl})$ z dopuszczalną (krytyczną) wartością α_{kr}.

W przypadku drugiej metody weryfikacji hipotezy zerowej dąży się do porównania poziomów istotności α, powyżej których odpowiednie wielkości wejściowe mają istotny wpływ na wielkość wyjściową. Wyznacza się wówczas poziom istotności $\alpha(F_{obl})$ dla znanych wartości F_{obl} i stopni swobody f_1 i f_2. Wyznaczona wartość $\alpha(F_{obl})$ upoważnia do stwierdzenia, że mamy do czynienia z istotnością wpływu czynnika wejściowego na wyjściowy dla każdej wartości przyjmowanego poziomu istotności $\alpha=\alpha_{kr}$ nie mniejszej od $\alpha(F_{obl})$ [Polański, 1984]:

$$\alpha = \alpha_{kr} \geq \alpha(F_{kr}) \tag{1}$$

Wyznaczenie poziomu istotności $\alpha(F_{obl})$ możliwe jest dzięki zastosowaniu tablic statystycznych lub z użyciem dostępnych kalkulatorów statystycznych poziomu istotności – rys. 2a [policzto.com.pl].

a) b)



Rys. 2. Kalkulatory statystyczne służące do wyznaczania poziomu istotności α_{kr} dla znanych wartości stopni swobody f_1 i f_2 i wartości F_{kr} (a) oraz do obliczania wartości F_{kr} dla znanych wartości stopni swobody f_1 i f_2 i założonego poziomu istotności α (b)

Dla znanych wartości stopni swobody f_1 i f_2 oraz przyjętego poziomu istotności α możliwe jest wyznaczenie wartości krytycznej statystyki F_{kr} rozkładu Fishera-Snedecora. Można to zrobić z użyciem kalkulatora statystycznego – rys. 2b [wolframalpha.com] albo posługując się odpowiednią zależnością (2) [Zieliński, 1972]. Zmienna losowa $F_{f1,f2}$ ma rozkład Fishera-Snedecora, gdy jej dystrybuanta rozkładu ciągłego wyraża się poniższym wzorem:

$$P(F_{f_1,f_2} < x) = \frac{\Gamma\left(\frac{f_1+f_2}{2}\right)}{\Gamma\left(\frac{f_1}{2}\right)\Gamma\left(\frac{f_2}{2}\right)} f_1^{\left(\frac{f_1}{2}\right)} f_2^{\left(\frac{f_2}{2}\right)} \int_0^x t^{\left(\frac{f_1}{2}-1\right)} (f_2 - f_1 t)^{\left(\frac{f_1+f_2}{2}\right)} dt \qquad x \geq 0 \tag{2}$$

gdzie: $F_{f1,f2}$ – rozkład zmiennej losowej, Γ – rozkład gamma, f_1 i f_2 – stopnie swobody odpowiednio dla licznika i mianownika, x – wartość zmiennej losowej, t – realizacja zmiennej losowej

Wartością krytyczną rozkładu Fishera-Snedecora nazywa się liczbę $F_{kr(\alpha,f1,f2)}$ spełniającą następujący warunek:

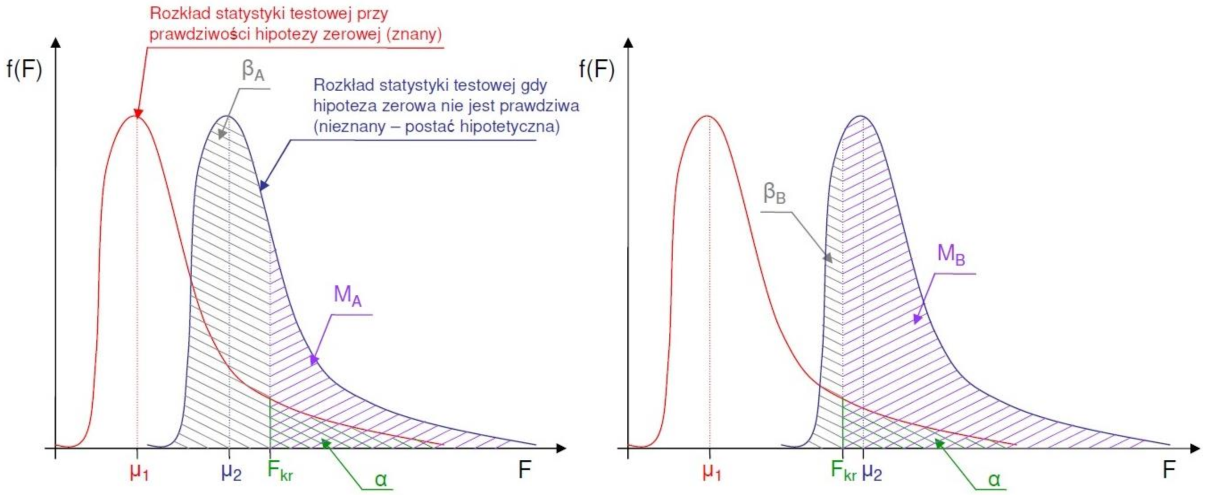
$$P(F_{f_1,f_2} \geq F_{kr(\alpha,f_1,f_2)}) = \alpha_{kr} \tag{3}$$

Istotnym zagadnieniem statystycznej weryfikacji hipotez jest moc testu [Smirnow i Dunin–Barkowski, 1969]. Jest on prawdopodobieństwem odrzucenia fałszywej hipotezy zerowej (niepopełnienia błędu drugiego rodzaju) [Wasilewska, 2015]. Moc testu M statystycznego wyraża się zależnością:

$$M=1-\beta \tag{4}$$

gdzie β jest prawdopodobieństwem popełnienia błędu II rodzaju.

Moc testu jest zatem miarą skuteczności testu do wykrywania fałszywej hipotezy zerowej. Im większa jest moc testu, tym lepszy (mocniejszy) jest zastosowany test. Większa moc oznacza, że test z większą skutecznością wykrywa fałszywą hipotezę zerową. Gdy test jest słaby istnieje duże prawdopodobieństwo przyjęcia hipotezy zerowej, gdy jest ona fałszywa. Moc testu zależy przede wszystkim od: poziomu istotności α, odchylenia standardowego w populacji σ, wielkości próby *n*, różnicy wartości średnich μ_1 i μ_2 (będących miarą położenia badanej cechy populacji). Wybierając odpowiedni obszar krytyczny należy to wykonać w taki sposób, aby weryfikacja sprawdzanej hipotezy miała jak największą czułość, czyli aby wartość M była jak największa. Oznacza to, że istnieje największe prawdopodobieństwo, że zmienna z próby przyjmie wartość z obszaru krytycznego, gdy prawdziwa jest hipoteza konkurencyjna do tej sprawdzanej – rys. 3.



Rys. 3. Interpretacja graficzna mocy testu statystycznego: α - poziom istotności, β - prawdopodobieństwo popełnienia błędu II rodzaju, M - moc testu, μ_1 - wartość średnia rozkładu statystyki testowej przy prawdziwości hipotezy zerowej, μ_2 - hipotetyczna wartość średnia rozkładu statystyki testowej przy nieprawdziwości hipotezy zerowej, indeksy na rysunku: A - niska moc statystyczna testu, B - wysoka moc statystyczna testu

Dodatkowo w analizach statystycznych zdefiniowano wartość współczynnika pewności istotności wpływu[2] ΔF jako miary mocy testu statystycznego, będącego różnicą pomiędzy wartością statystyki F_{obl}, obliczoną dla badanego czynnika wejściowego, z wartością krytyczną dla założonych wartości stopni swobody licznika i mianownika oraz poziomu istotności:

$$\Delta F=F_{kr}-F_{obl} \tag{5}$$

Wartość obliczeniowa statystyki F_{obl} jest definiowana jako stosunek wariancji wielkości wejściowej do wariancji charakteryzującej niedokładność pomiarów. Zatem im większe jest odchylenie od wartości średniej czynnika wejściowego, tym większa jest wartość F_{obl}. W tym przypadku, przy określonej wartości krytycznej F_{kr}, współczynnik pewności ΔF rośnie wraz ze wzrostem F_{obl}. Konkludując, im większe wartość przyjmuje ΔF, tym większy jest wpływ

[2] Dalej nazywanego krócej "współczynnikiem pewności"

czynnika wejściowego na wyjściowy, zatem wartość współczynnika pewności ΔF pozwala na ocenę jak istotny jest ten wpływ [Smirnow i Dunin–Barkowski, 1969].

Najwygodniejsze jest zastosowanie analizy współczynnika pewności dla pierwszego sposobu weryfikacji H_0 (dzięki porównaniu wartości statystyki: obliczeniowej F_{obl} i krytycznej F_{kr}). W przypadku drugiej metody weryfikacji hipotezy zerowej, gdzie porównuje się wartości poziomu istotności α_{kr}, powyżej których wielkość wejściowa ma istotny wpływ na wyjściową, analiza współczynnika pewności ΔF jest również możliwa. Wyznaczając $\alpha(F_{obl})$, dla której mamy do czynienia z istotnością wpływu, zakładamy że współczynnik pewności $\Delta F=F_{kr}–F_{obl}=0$, ponieważ $F_{kr}=F_{obl}$ dla zmiennych wartości $\alpha(F_{obl})$. Wówczas coraz większe wartości F_{obl} (coraz większy wpływ czynnika wejściowego na wyjściowy) skutkują coraz mniejszymi wartościami poziomu istotności $\alpha(F_{obl})$. Istnieje wówczas coraz mniejsze prawdopodobieństwo odrzucenia prawdziwej H_0. Jednakże należy pamiętać, że zmniejszanie wartości α jednocześnie zwiększa prawdopodobieństwo popełnienia błędu II rodzaju, wynoszące $1–\alpha$.

3. Wpływ wartości mocy testu na wyniki analizy statystycznej

Przeanalizowano wpływ wartości poziomu istotności α na wartość krytyczną statystyki F_{kr} i oceniono, czy zwiększanie mocy testu statystycznego M (poprzez zwiększanie wartości α przy jednoczesnym zmniejszaniu wartości β) jest uzasadnione [Jarmakowska-Kostrzanowska, 2021]. Sformułowano następujące hipotezy:

Hipoteza zerowa **H_0**: obciążenie silnika <u>ma wpływ</u> na wartość entalpii spalin wylotowych
Hipoteza alternatywna **H_A**: obciążenie silnika <u>nie ma wpływu</u> na wartość entalpii spalin wylotowych

Należy zauważyć, że zwiększając wartość α zwiększa się prawdopodobieństwo odrzucenia H_0, gdy jest ona prawdziwa oraz przyjęcia H_A (w tym przypadku fałszywej). Jednocześnie zmniejsza się wartość β, zatem zmniejsza się prawdopodobieństwo przyjęcia fałszywej H_0 i odrzucenia H_A jako fałszywej. Zazwyczaj konsekwencje przyjęcia hipotezy fałszywej są znacznie poważniejsze i bardziej szkodliwe niż skutki odrzucenia hipotezy prawdziwej [Korzyński, 2017].

Tab. 2. Wartość poziomu istotności α w zależności od zadanej wartości krytycznej statystyki F_{kr}, dla liczby stopni swobody: licznika p=3 i mianownika q=3, wyznaczona w kalkulatorze statystycznym [policzto.com.pl]

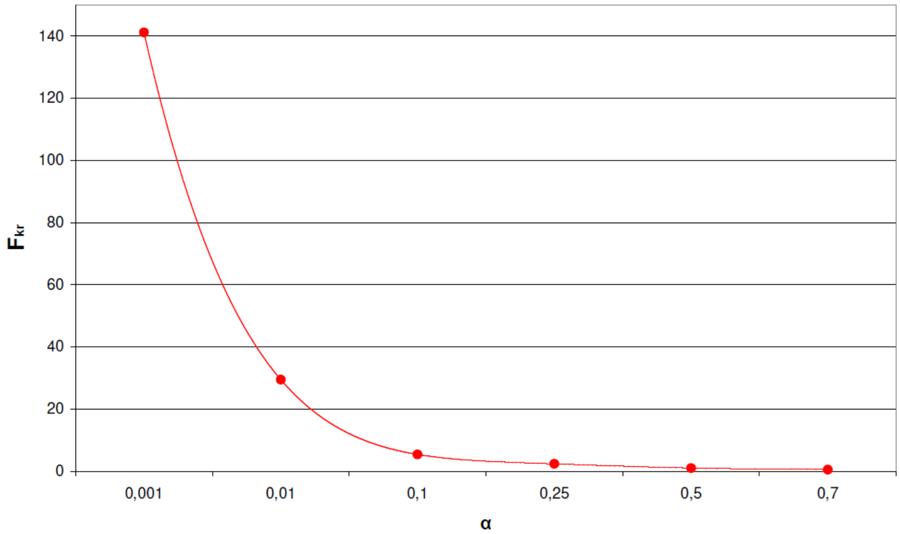
Wartość krytyczna statystyki F_{kr}	Poziom istotności α
0,5	0,7082
1	0,5
5	0,1096
10	0,0452
50	0,0046
100	0,0017

Tab. 3. Wartość krytyczna statystyki F_{kr} w zależności od zadanej wartości poziomu istotności α, dla liczby stopni swobody: licznika p=3 i mianownika q=3, wyznaczona w kalkulatorze statystycznym [wolframalpha.com.pl]

Poziom istotności α	Wartość krytyczna statystyki F_{kr}
0,001	141,108
0,01	29,457
0,1	5,39
0,25	2,355
0,5	1
0,7	0,5155

Przeprowadzono analizę w kalkulatorze statystycznym na stronie internetowej [policzto.com.pl], podczas której zwiększano wartość F_{kr} i uzyskiwano wartości α - tab. 2. Pozwoliło to ocenić czy zwiększanie F_{kr} (zależnego przede wszystkim od α i liczby stopni swobody licznika i mianownika, które ustalono jako stałe) ma uzasadnienie merytoryczne [Puzdrowska, 2021b]. Wykonano podobną analizę w programie komputerowym na stronie [wolframalpha.com], gdzie przeanalizowano jak zmniejszanie wartości α wpływa na F_{kr}, przy tej samej liczbie stopni swobody licznika i mianownika - tab. 3. Pozwoliło to na ocenę czy dążenie, aby wartość poziomu istotności α była jak najmniejsza jest celowe.

Na rysunku 4 przedstawiono w postaci graficznej zależność $F_{kr}=f(\alpha)$ na podstawie danych z tabeli 3. Jak widać na rysunku zmniejszanie wartości poziomu istotności α skutkuje wzrostem wartości krytycznej statystyki F_{kr}, co jest zbieżne z interpretacją graficzną tych dwóch wartości przedstawioną na rys. 1. Należy jednak zwrócić uwagę, że obniżenie wartości α poniżej 0,01 powoduje znacznie szybszy wzrost wartości F_{kr}. Przy tak małych wartościach poziomu istotności spada prawdopodobieństwo odrzucenia prawdziwej hipotezy zerowej, jednakże rośnie możliwość popełnienia błędu drugiego rodzaju β, a co za tym idzie spada moc testu statystycznego - test wykrywa fałszywą hipotezę zerową z mniejszą skutecznością - rys. 3.



Rys. 4. Wpływ wartości poziomu istotności α na wartości krytyczne statystyki F_{kr}, dla liczby stopni swobody: licznika p=3 i mianownika q=3

4. Wnioski i podsumowanie

Z przeprowadzonych analiz w kalkulatorach statystycznych wynika, że zbytnie zmniejszanie wartości poziomu istotności α (lub zwiększanie wartości krytycznej F_{kr}) rozkładu Fishera-Snedecora) jest nieuzasadnione. Zmniejszanie wartości α poniżej 0,01 skutkuje spadkiem mocy testu statystycznego M oraz wzrostem prawdopodobieństwa popełnienia błędu drugiego rodzaju β, przy bardzo małym prawdopodobieństwie odrzucenia prawdziwej hipotezy zerowej. Zatem, opracowując metodykę diagnozowania elementów ograniczających komorę spalania okrętowego silnika o ZS wraz z aparaturą paliwową oraz kanałami powietrza dolotowego i spalin wylotowych na podstawie szybkozmiennej temperatury spalin wylotowych i stosując w analizie test F rozkładu Fishera-Snedecora należy tak dobrać poziom istotności α oraz wartość krytyczną F_{kr}, aby nie ucierpiały na tym wartości β i M. Wartość mocy testu statystycznego M również powinna być optymalna tak, aby interpretacja postawionej hipotezy zerowej była odpowiednia [Jarmakowska-Kostrzanowska, 2021]. Zachowanie "złotego środka" w doborze parametrów testu statystycznego może pozwolić na uzyskanie jak największej jego skuteczności, zatem i na jak najtrafniejsze wnioskowanie merytoryczne na podstawie jego wyników.

Bibliografia

[1] Jarmakowska-Kostrzanowska, L. (2021), Istotnie statystyczna moc testu – analiza mocy i jej miejsce w przyborniku badacza oraz interpretacja (nie)istotności statystycznej przy małej (dużej) mocy testu. Przegląd Psychologiczny, 64(4), 83–99. https://doi.org/10.31648/pp.7359
[2] Korczewski Z., Puzdrowska P. (2015), Analytical method of determining dynamic properties of thermocouples used in measurements of quick – changing temperatures of exhaust gases in marine diesel engines. Poznań. Combustion Engines, nr 162 (3) (2015), s. 300-306.
[3] Korzyński M. (2017), Metodyka eksperymentu. Planowanie, realizacja i statystyczne opracowanie wyników eksperymentów technologicznych. Wydawnictwo Naukowo Techniczne. Warszawa.
[4] Pabis S. (1985), Metodologia i metody nauk empirycznych. Państwowe Wydawnictwo Naukowe. Warszawa.
[5] Piotrowski I. (1976), Podstawy metrologii. Państwowe Wydawnictwo Naukowe.
[6] Polański Z. (1984), Planowanie doświadczeń w technice. Państwowe Wydawnictwo Naukowe. Warszawa.
[7] Puzdrowska P. (2018), Signal filtering method of the fast-varying diesel exhaust gas temperature. Combustion Engines, nr. 175(4), s.48-52.
[8] Puzdrowska P. (2019), Statystyka F rozkładu Fishera-Snedecora jako narzędzie do oceny istotności wpływu mocy silnika o ZS na wybrane miary diagnostyczne. Journal of Polish CIMAC. -Vol. 14, nr. 1/18, s. 177-186.
[9] Puzdrowska P. (2021a), Application of the F-statistic of the Fisher-Snedecor distribution to analyze the significance of the effect of changes in the compression ratio of a diesel engine on the value of the specific enthalpy of the exhaust gas flow. Combustion Engines, *186*, 80-88.
[10] Puzdrowska P.(2021b), Diagnostic information analysis of quickly changing temperature of exhaust gas from marine diesel engine part i single factor analysis. Polish Maritime Research -Vol. 28,iss. 4(112) (2021), s.97-106.
[11] Smirnow N.W., Dunin-Barkowski I.W. (1969), Kurs rachunku prawdopodobieństwa i statystyki matematycznej dla zastosowań technicznych. PWN. Warszawa.
[12] Wasilewska E. (2015), Statystyka matematyczna w praktyce. Wydawnictwo Difin. Warszawa.
[13] Zieliński R. (1972), Tablice statystyczne. Państwowe Wydawnictwo Naukowe.
[14] *www.policzto.com.pl* dostęp 10.07.2022
[15] *www.wolframalpha.com* dostęp 10.07.2022