Olimpia Pozauć, Oliwia Krzemińska

Sprawozdanie 2
grupa 12

23 stycznia 2022

Spis treści

1. Wprowadzenie

Celem sprawozdania jest analiza trzech testów statystycznych i sprawdzenie przypadków, w których założenia testów nie zostały spełnione. Przyjrzymy się sytuacjom, w których dane będą pochodzić z rozkładu normalnego oraz wykładniczego. Za każdym razem postaramy się ustalić, który spośród testów statystycznych jest jednoznacznie najmocniejszy na poziomie istotności α.W sprawozdaniu będziemy rozważać testy na poziomie istotności $\alpha = 0.05$ do testowania:
— $H_0 : \mu = 3$, przeciwko
— $H_1 : \mu < 3$.
Będziemy wykonywać sprawdzenie tej hipotezy, stosując wprost trzy testy:
1. test z (przy założeniu $\sigma = 2$),
2. test t-Studenta,
3. test rang znakowanych Wilcoxona.

1.1. Definicje i oznaczenia

— $\mathbf{H_0} : \mathbf{\Theta} \in \mathbf{\Theta_0}$ - Hipoteza zerowa.
— $\mathbf{H_1} : \mathbf{\Theta} \in \mathbf{\Theta_1}$ - Hipoteza alternatywna.
— **Funkcja mocy testu**
$\beta_\varphi : \Theta \to [0, 1]$, to

$$\beta_\varphi(\Theta) = \mathbb{E}_\Theta[\varphi(X)],$$

gdzie $\mathbb{E}$ to wartość oczekiwana, φ to funkcja mocy testu, a X to próba losowa. Dla testów niezrandomizowanych, czyli takich, dla których φ przyjmuje jedynie wartośći 0 i 1, możemy zapisać

$$\beta_\varphi(\Theta) = P_\Theta[\varphi(X) = 1],$$

czyli jest to prawdopodobieństwo odrzucenia hipotezy zerowej.
— **Moc testu** - wartość funkcji mocy w punktach $\Theta \in \Theta_1$ przy alternatywie Θ.
— **Rozmiar testu** - kres górny odrzucenia prawdopodobieństwa, gdy prawdziwy parametr należy do Θ_0. Jest górnym ograniczeniem popełnienia błędu I rodzaju, czyli odrzucenia prawdziwej hipotezy zerowej i zapisujemy go jako

$$\beta_\varphi = \sup_{\theta \in \Theta_0} \beta_\varphi(\theta).$$

— **Poziom istotności testu** - ograniczenie rozmiaru testu przez $\alpha \in (0, 1)$.
— **Test jednostajnie najmocniejszy na poziomie istotności** α - test na poziomie istotności α, którego wartość funkcji mocy jest przynajmniej taka sama jak wartość funkcji mocy innych testów na poziomie istotności α.

1.2. Opis wykorzystanych testów

1.2.1. Test z

Test z to test statystyczny, dla którego rozkład statystyki testowej przy hipotezie zerowej może być przybliżony rozkładem normalnym. Testy z badają średnią rozkładu oraz dla każdego poziomu istotności w przedziale ufności mają jedną wartość krytyczną. Dla próby $(X_1, X_2, .., X_n) \sim \mathcal{N}(\mu, \sigma^2)$, gdzie σ jest znana. Możemy zapisać ogólne postacie testów:

— $H_0 : \mu = \mu_0$

$$\overline{X} \sim \mathcal{N}(\mu_0, \sigma^2),$$

— $H_1 : \mu < \mu_0$

$$\varphi(x) = \begin{cases} 1, \text{dla } \overline{X} \leqslant c, \\ 0, \text{dla } \overline{X} > c, \end{cases}$$

gdzie $c \in \mathbb{R}$. Poziom istotności α wygląda następująco :

$$\alpha = P_{\mu_0}(\sqrt{n} \cdot \frac{\overline{X} - \mu_0}{\sigma} < \sqrt{n} \cdot \frac{c - \mu_0}{\sigma}).$$

Natomiast funkcja mocy prezentuje się:

$$\beta_\varphi(\mu) = P_\mu(\overline{X} \leqslant \mu_0 - z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}).$$

1.2.2. Test t-Studenta

Test t jest dowolnym testem hipotezy statystycznej, w którym statystyka testowa jest zgodna z rozkładem t-Studenta przy hipotezie zerowej. Możemy zapisać:

$$t = \sqrt{n}\frac{\overline{X} - \mu_0}{S} \sim \tau(n-1),$$

gdzie

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}.$$

Zapiszemy teraz ogólną postać testu oraz poziom istotności α zakładając przy tym, że $c \in \mathbb{R}$:
— $H_0 : \mu = \mu_0$ oraz $H_1 : \mu < \mu_0$:

$$\varphi(x) = \begin{cases} 1, \text{dla } \sqrt{n}\frac{\overline{X}-\mu_0}{S} \geqslant t_{n-1}(1-\alpha), \\ 0, \text{dla } \sqrt{n}\frac{\overline{X}-\mu_0}{S} > t_{n-1}(1-\alpha). \end{cases}$$

Poziom istotności możemy zapisać jako:

$$\alpha = P_{\mu_0}(\varphi(X) = 1) = P_{\mu_0}(t \geqslant c) = 1 - P_{\mu_0(t \leqslant c)}.$$

Natomiast funkcja mocy ma postać:

$$\beta_\varphi(X) = P_\mu(\sqrt{n}\frac{\overline{X} - \mu_0}{S} \leqslant t_{n-1}(\alpha)) = 1 - F_{n-1;\sqrt{n}\frac{\mu-\mu_0}{\sigma}}(t_{n-1}(1-\alpha)).$$

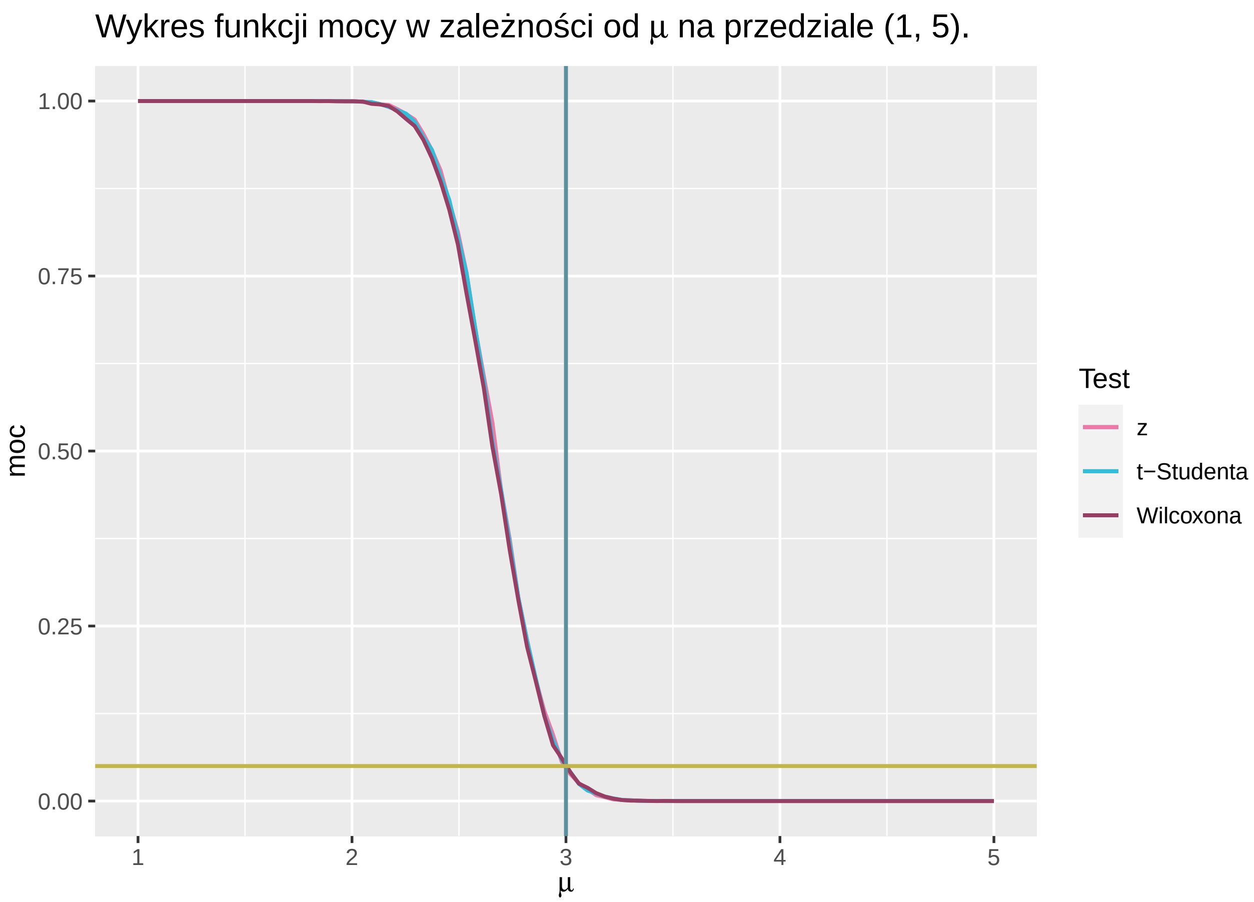
1.2.3. Test rang znakowanych Wilcoxona

Testem rangowym nazywamy test, w którym statystyka testowa jest konstruowana w oparciu o rangi. Test Wilcoxona jest często stosowanym nieparametrycznym odpowiednikiem testu t-Studenta. W założeniu nie służy on w ogólności do testowania średniej, jednak możemy go używać, kiedy mamy dozynienia z rozkładami symetrycznymi, w których mediana jest równa.

2. Zadanie 1

2.1. Treść

Rozważmy próbę $(X_1, ..., X_{100})$ z rozkładu normalnego $\mathcal{N} \sim (\mu, 2^2)$. Korzystając z symulacji Monte Carlo wykonaj wykres funkcji mocy w zależności od μ na przedziale $(1, 5)$ dla wszystkich trzech testów. Czy istnieje test jednostajnie najmocniejszy spośród nich?



Rysunek 1. Wykres funkcji mocy testów w zależności od μ dla próby z rozkładu $\mathcal{N} \sim (\mu, 2^2)$ o długości $n = 100$ przy 10^4 krokach Monte Carlo dla $H_0 : \mu = 3$ przeciwko $H_1 : \mu < 3$.

2.2. Przedstawienie funkcji mocy na wykresie

Na wykresie przedstawimy funkcję mocy testu dla trzech testów: testu z (przy założeniu $\sigma = 2$), testu t-Studenta oraz testu rang znakowanych Wilcoxona w zależności od μ na przedziale $(1, 5)$ dla danych z rozkładu normalnego $\mathcal{N} \sim (\mu, 2^2)$.

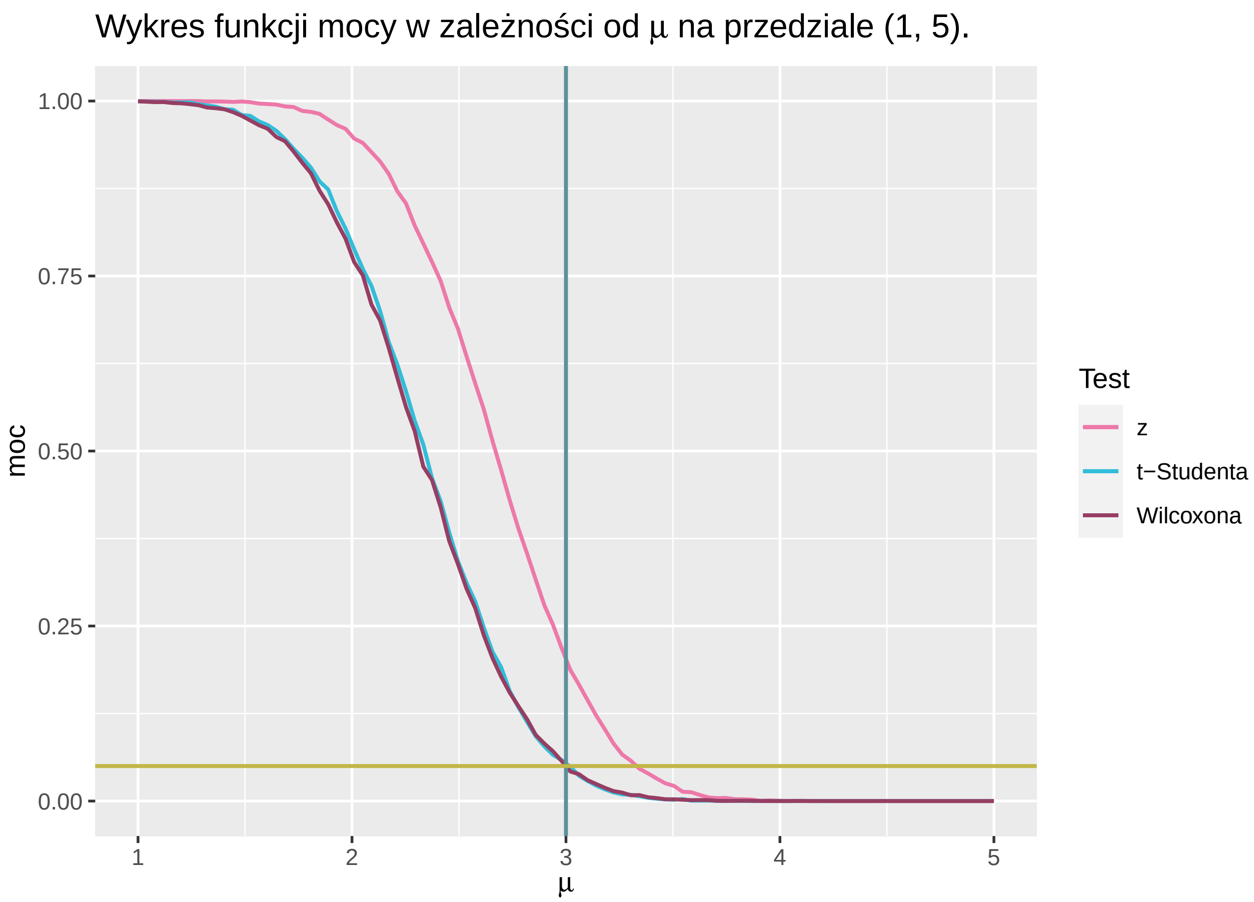
2.3. Wnioski

Na podztawie wykresu 1 możemy zauważyć, że wszystkie trzy testy, czyli test z, test t-Studenta oraz test Wilcoxona są nieobciążone, ponieważ dla hipotezy alternatywnej $H_1 : \mu < 3$ moc testu jest większa od zadanego poziomu istotności $\alpha = 0.05$.

W naszym przypadku jest jeden punkt hipotezy zerowej $H_0 = 3$. Wartość mocy wszystkich testów dla tego punktu jest równa w przybliżeniu 0.05, zatem możemy wnioskować, że dla danych przedstwionych na wykresie 1 rozmiar testów to około 0.05.

Dla wszystkich testów możemy również przyjąć, że ich rozmiar jest ograniczony przez α, zatem są one na poziomie istotności $\alpha = 0.05$.

Nie możemy więc jednoznacznie określić, który test jest jednostajnie najmocniejszy, ponieważ nie możemy dostrzec znaczących różnic między nimi.



Rysunek 2. Wykres funkcji mocy testów w zależności od μ dla próby z rozkładu $\mathcal{N} \sim (\mu, 4^2)$ o długości $n = 100$ przy 10^4 krokach Monte Carlo dla $H_0 : \mu = 3$ przeciwko $H_1 : \mu < 3$.

3. Zadanie 2

3.1. Treść

Rozważmy próbę $(X_1, ..., X_{100})$ z rozkładu normalnego $\mathcal{N} \sim (\mu, 4^2)$. Wykonaj wykres funkcji mocy na wybranym przedziale zawierającym przynajmniej po jednym punkcie z hipotezy zerowej i alternatywnej. Jak zmieniła się funkcja mocy testów? Czy w tym przypadku istnieje test jednostajnie najmocniejszy spośród nich?

3.2. Przedstawienie funkcji mocy na wykresie

Na wykresie ponownie przedstawimy funkcję mocy testu dla trzech testów: testu z (przy założeniu $\sigma = 2$), testu t-Studenta oraz testu rang znakowanych Wilcoxona w zależności od μ. Wybrany przez nas przedział to ponownie $(1, 5)$. Tym razem dane pochodzą z rozkładu normalnego $\mathcal{N} \sim (\mu, 4^2)$.

3.3. Wnioski

Na podstawie wykresu 2 zauważamy, że dla testu t-Studenta oraz Wilcoxona możemy wyciągnąć takie same wnioski, jak w zadaniu 1. Oba te testy są nieobciążone, rozmiar testów wynosi około 0.05 i zachowują one poziom istotności $\alpha = 0.05$.

Nie możemy zatem wybrać z tych dwóch testów jednostajnie mocniejszego, ponieważ ponownie nie dostrzegamy znaczących różnic pomiędzy nimi.

Natomiast test z wypadł na wykresie 2 znacząco inaczej niż poprzednio. Nie zachowuje on zadanego poziomu istotności $\alpha = 0.05$. Jest to spowodowane zignorowaniem założeń, które należało spełnić. Przyjeliśmy $\sigma = 2$, kiedy należało przyjąć $\sigma = 4$, zgodnie z zadanym rozkładem.

W związku z tym nie bierzemy go pod uwagę w roztrzyganiu, który z testów jest jednostajnie najmocniejszy.

4. Zadanie 3

4.1. Treść

Rozważmy próbę $(X_1, ..., X_{100})$ z rozkładu wykładniczego $\mathcal{E} \sim \left(\frac{1}{\mu}\right)$. Wykonaj wykres funkcji mocy na wybranym przedziale zawierającym przynajmniej po jednym punkcie z hipotezy zerowej i alternatywnej. Jak zmieniła się funkcja mocy testów? Czy w tym przypadku istnieje test jednostajnie najmocniejszy spośród nich?

4.2. Przedstawienie funkcji mocy na wykresie

Na wykresie kolejny raz przedstawimy funkcję mocy testu dla trzech testów: testu z (przy założeniu $\sigma = 2$), testu t-Studenta oraz testu rang znakowanych Wilcoxona w zależności od μ. Wybrany przez nas przedział to ponownie $(1, 5)$. Tym razem dane pochodzą z rozkładu wykładniczego $\mathcal{E} \sim \left(\frac{1}{\mu}\right)$.

Ponieważ dane pochodzą z rozkładu wykładniczego, a jak wiemy rozkład ten nie jest symetryczny, dlatego test Wilcoxona nie zadziała poprawnie. W związku z tym użyliśmy skorygowanego wariantu, w którym traktujemy test jako test pseudomedian.

4.3. Wnioski

Na wykresie 3 widzimy, że w przypadku, gdy dane pochodzą z rozkładu wykładniczego tylko test Wilcoxsona zachowuje poziom istotności $\alpha = 0.05$, a jego rozmiar w przybliżeniu również wynosi 0.05.

Ponieważ tylko test Wilcoxona spełnia wszystkie wymagania, możemy jednoznacznie stwierdzić, że jest on w tym przypadku jednostajnie najmocniejszy.

Co ciekawe, na wykresie 3 możemy zauważyć także, że test t-Studenta jest bardzo blisko spełnienia wymagań, zatem możemy stwierdzić, że jest on mocniejszy niż test z.

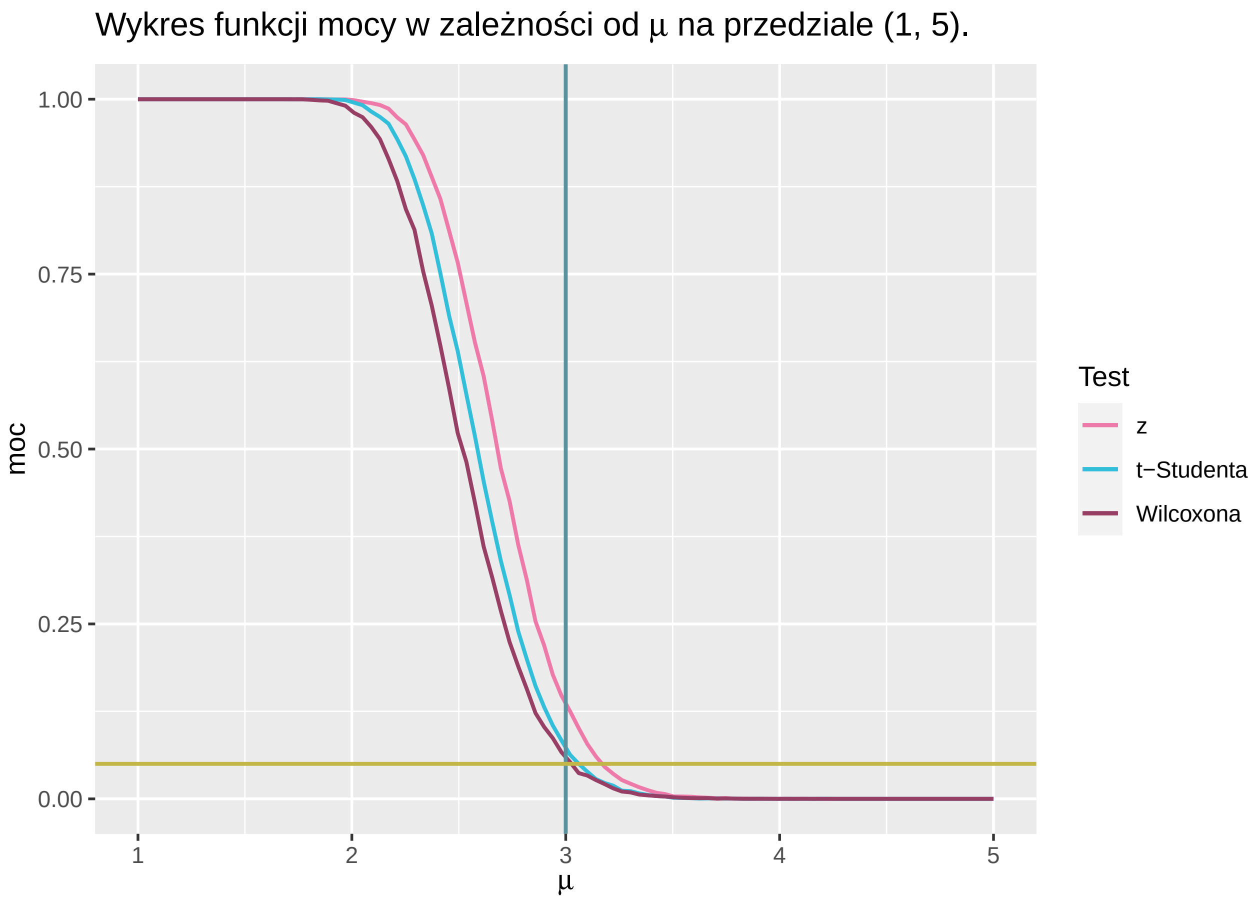
5. Podsumowanie

Celem sprawozdania była anliza mocy trzech testów statystycznych oraz aprwdzenie przypadków, w których ich założenia nie zostały spełnione.

Braliśmy pod uwagę trzy testy: test z, test t-Studenta oraz test rang znakowanych Wilcoxona w trzech różnych sytuacjach.

Na podstawie przeprowadzonej analizy możemy stwierdzić, że najlepiej wypadł test Wilcoxona, ponieważ za każdym razem zadziałał on poprawnie, a poziom istotności $\alpha = 0.05$ został zachowany.

Warto pamiętać, że działa on jednak poprawnie, tylko dla rozkładów symetrycznych, w innym przypadku należy dokonać korekcji.



Rysunek 3. Wykres funkcji mocy testów w zależności od μ dla próby z rozkładu $\mathcal{E} \sim \left(\frac{1}{\mu}\right)$ o długości $n = 100$ przy 10^4 krokach Monte Carlo dla $H_0 : \mu = 3$ przeciwko $H_1 : \mu < 3$.

Możemy również zaobserwować, że w przypadku rozkładu wykładniczego testy nie działają poprawnie. Związane jest to z niespełnieniem założeń, które mówią o tym, że próba powinna być z rozkładu normalnego. Podobne wnioski możemy wyciągnąć z zadania drugiego, gdzie nie dotrzymaliśmy założeń testu z, wybierając inną σ.

Warto pamiętać, że aby testy działały dobrze należy spełnić wszystkie założenia.

6. Źródła

1. https://en.wikipedia.org/wiki/Z-test

2. http://prac.im.pwr.edu.pl/~giniew/lib/exe/fetch.php?media=rok2122:zimowy:07.pdf
3. http://prac.im.pwr.edu.pl/~giniew/lib/exe/fetch.php?media=rok2122:zimowy:08.pdf
4. http://prac.im.pwr.edu.pl/~giniew/lib/exe/fetch.php?media=rok2122:zimowy:09.pdf
5. http://prac.im.pwr.edu.pl/~giniew/lib/exe/fetch.php?media=rok2122:zimowy:10.pdf